Comparison of Cloud Microphysics Schemes in a Warn-on-Forecast System Using Synthetic Satellite Objects

Thomas A. Jones, Patrick Skinner, and Kent Knopfmeier

Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

Edward Mansell

National Severe Storms Laboratory, Norman, Oklahoma

Patrick Minnis

NASA Langley Research Center, and Science Systems and Applications, Inc., Hampton, Virginia

Rabindra Palikonda

Science Systems and Applications, Inc., Hampton, Virginia

William Smith Jr.

NASA Langley Research Center, Hampton, Virginia

ABSTRACT

Forecasts of high-impact weather conditions using convection-allowing numerical weather prediction models have been found to be highly sensitive to the selection of cloud microphysics scheme used within the system. The Warn-on-Forecast (WoF) project has developed a rapid-cycling, convection-allowing, data assimilation and forecasting system known as the NSSL Experimental WoF System for ensembles (NEWS-e), which is designed to utilize advanced cloud microphysics schemes. NEWS-e currently (2017–18) uses the double-moment NSSL variable density scheme (NVD), which has been shown to generate realistic representations of convective precipitation within the system. However, very little verification on nonprecipitating cloud features has been performed with this system. During the 2017 Hazardous Weather Testbed (HWT) experiment, an overestimation of the areal coverage of convectively generated cirrus clouds was observed. Changing the cloud microphysics scheme to Thompson generated more accurate cloud fields. This research undertook the task of improving the cloud analysis generated by NVD while maintaining its skill for other variables such as reflectivity. Adjustments to cloud condensation nuclei (CCN), fall speed, and collection efficiencies were made and tested over a set of six severe weather cases occurring during May 2017. This research uses an object-based verification approach in which objects of cold infrared brightness temperatures, high cloud-top pressures, and cloud water path are generated from model output and compared against *GOES-13* observations. Results show that the modified NVD scheme generated much more skillful forecasts of cloud objects than the original formulation without having a negative impact on the skill of simulated composite reflectivity forecasts.

1. Introduction

Synthetic satellite imagery has been used to visualize and verify numerical weather prediction (NWP) model output for over two decades. In particular, a substantial amount of research has focused on simulated geostationary imager water vapor and infrared radiances to track cloud coverage, determine convective initiation, and relate cloud properties to potential high-impact weather events (e.g., Chevallier et al. 2001; Chevallier and Kelly 2002; Grasso and Greenwald 2004; Chaboureau and Pinty 2006; Mecikalski and Bedka 2006; Grasso et al. 2008; Cintineo et al. 2013, 2014; Griffin et al. 2017a,b).

Corresponding author: Dr. Thomas A. Jones, thomas.jones@noaa.gov

Simulated satellite radiances are calculated by ingesting forecast model atmospheric and surface conditions into a radiative transfer model (RTM) that relates these conditions to expected satellite observations based on satellite sensor lookup tables generated for individual wavelengths and fields of view (Weng 2007; Han et al. 2007). Using this method, simulated satellite imagery has become an indispensable tool in current forecasting environments and is particularly useful during high-impact weather events (e.g., Line et al. 2016; Lindsey et al. 2018). For example, synthetic infrared radiances can be used to assess the depth and intensity of forecast convection by measuring the temperature of cloud tops and where convection is developing prior to the onset of precipitation. Other infrared radiances are sensitive to mid- and upper-level atmospheric moisture and can be used to assess a modeled environment's favorability for high-impact weather to occur (Jones et al. 2018). The location and thickness of clouds has a direct impact on solar radiation reaching the surface, resulting in modification to the near-surface thermodynamic conditions (e.g., Xie et al. 2012; Cintineo et al. 2014; Jones et al. 2015). As a result, synthetic satellite imagery can be used to determine where these impacts are likely to be maximized.

In addition to being an important forecast tool, synthetic satellite imagery has been used for model verification (e.g., Tselioudis and Jakob 2002; Keil et al. 2003; Otkin et al. 2009; Grasso et al. 2010; Matsui et al. 2014; Griffin et al. 2017a,b). Additional studies focus on verification of cloud properties and how different cloud model microphysics characterize them (Grasso and Greenwald 2004; Liu and Moncrieff 2007; Chaboureau and Pinty 2006; Otkin and Greenwald 2008; Grasso and Lindsey 2011; Cintineo et al. 2014; Grasso et al. 2014). Cloud microphysics schemes are included in NWP models to represent the formation of, interaction between, and dissipation of liquid and frozen hydrometeors. Each scheme contains different assumptions on the properties of cloud and precipitation hydrometeors, which can be a significant source of model error (e.g., Snook and Xue 2008; Tong and Xue 2008). Single-moment microphysics schemes only predict hydrometeor mixing ratios while more complex double-moment schemes predict both mixing ratios and hydrometeor number concentrations. Chaboureau and Pinty (2006) and Liu and Moncrieff (2007) found that the choice of cloud microphysics schemes had the greatest impact on upper-tropospheric hydrometeor concentrations. Otkin and Greenwald (2008) noted that the more complex schemes generally produced superior representations of cloud properties. Finally, Cintineo et al. (2014) found that double-moment schemes such as Morrison (Morrison et al. 2009) and Milbrandt–Yau (Milbrandt and Yau 2005a,b) produced too much upper-level cloud cover, whereas other schemes such as Thompson (Thompson et al. 2004, 2008) and WRF double-moment 6-class (WDM6) produced more realistic representations. Matsui et al. (2014) combined microwave and infrared radiances from polar-orbiting satellite to further classify observed clouds into convection, stratus, or anvil types. However, since high temporal resolution microwave radiances are not available, we are unable to verify individual cloud types using those methods.

Verifying cloud microphysics schemes present within convection-permitting models using satellite data complements many studies using radar reflectivity (REFL) for the same purpose. For example, Dawson et al. (2012), Jung et al. (2012), and Yussouf and Stensrud (2012) noted that double-moment microphysics schemes generate a better representation of reflectivity within the model compared to single-moment schemes. This research extends these previous works by verifying several sets of experiments using *both* radar reflectivity and satellite imagery. In convection-permitting models, it is important to get both the characteristics of severe-weather-producing convection and the near-storm environment correct to generate accurate short-term (0–3 h) forecasts.

This research uses forecasts generated for high-impact weather events occurring in May 2017 using an ensemble data assimilation and forecasting system developed as part of the Warn-on-Forecast project (WoF; Stensrud et al. 2009; 2013), known as the NSSL Experimental WoF System for ensembles (NEWS-e; Wheatley et al. 2015; Jones et al. 2016). This system was run during the 2017 Hazardous Weather Testbed (HWT) in real time and used by forecasters to produce accurate short-term outlooks of convective storm hazards (Gallo et al. 2017; Choate et al. 2018). The 2017 version of the system used the NSSL full double-moment microphysics scheme (Mansell et al. 2010; Ziegler 1985) with variable-density graupel and hail (NVD). Using this configuration, a positive bias in the depth and coverage of cirrus cloud emanating from thunderstorm anvils was noted compared to the 2016 version of this system, which used Thompson microphysics (Thompson et al. 2004, 2008, 2016). To assess the impact of different microphysics on this system in a quantitative manner, six 2017 cases (9, 16, 17, 18, 23, and 27 May) were rerun using Thompson cloud microphysics and a modified NSSL double-moment scheme in order to reduce forecast biases related to cloudiness for future versions of the NEWS-e system. An object-based verification system (Skinner et al. 2016; 2018), based on the Method for Object Based Diagnostic Evaluation (MODE; Davis et al. 2006a,b), is used to identify both radar reflectivity and satellite-imagery-based objects to compare the various model

configurations. Each event considered was associated with multiple severe weather reports occurring during the late afternoon and evening within the region domain of each case.

Section 2 of this paper describes the NEWS-e system, microphysics schemes, and synthetic satellite products. Section 3 provides a description of the object verification methods and definitions on how radar reflectivity and satellite objects are defined. Results of the comparison of microphysics schemes are provided in section 4 with concluding remarks are presented in section 5.

2. Model configuration

a. 2017 Warn-on-Forecast system

A complete description of the NEWS-e system is provided in Wheatley et al. (2015), Jones et al. (2016), and Skinner et al. (2018) and is briefly summarized here. The configuration used for these experiments reflects the one used for HWT operational testing during the spring of 2017. The NEWS-e used the Advanced Research version of the Weather Research and Forecasting Model (WRF-ARW), version 3.8.1 (Skamarock et al. 2008). The data assimilation method is the parallel ensemble adjustment Kalman filter present in the Data Assimilation Research Testbed (DART) software (Anderson and Collins 2007; Anderson et al. 2009). Surface, radar, and satellite observations are assimilated into a 36-member ensemble at 15-min intervals starting at 1800 UTC each day and ending at 0300 UTC the next day. An experimental 36-member High Resolution Rapid Refresh ensemble (HRRRE) provides initial and boundary conditions for the 2017 version of the NEWS-e (Benjamin et al. 2016; Alexander et al. 2018). Both the HRRRE and NEWS-e systems use a 3-km horizontal grid spacing with 51 vertical levels and a model top at 20 hPa. The NEWS-e domain is 250 × 250 grid points or approximately 750 km × 750 km and is centered on the area of expected severe weather for each day. To maintain ensemble spread, each member uses a different set of boundary layer and radiation schemes (e.g., Stensrud et al. 2000; see Table 2 in Wheatley et al. 2015). One change from 2016 to 2017 was to replace the Thompson cloud microphysics scheme used in all members, with the NSSL double-moment scheme to potentially reduce overforecast biases observed in simulated reflectivity during the 2016 experiment (Skinner et al. 2018). This model configuration, while not perfect, has proven very successful during real-time testing and has been used in operational tornado and flash-flooding warning guidance. Additional improvements to the model configuration such as increased resolution and additional development in

microphysics will be applied to future versions of this system.

Assimilated conventional observations include surface temperature, humidity, wind, and pressure measurements from available Automated Surface Observing System (ASOS) sites and, if available, Oklahoma Mesonet sites within the NEWS-e domain. WSR-88D reflectivity contained within the 1-km Multi-Radar Multi-Sensor (MRMS) products and are objectively analyzed to 5-km resolution using a Cressman interpolation scheme (Cressman 1959; Smith et al. 2016). Radial velocity is processed directly from the level-II WSR-88D data, dealiased, and also objectively analyzed to a 5-km resolution using the same Cressman scheme. Satellite data in the form of cloud water path (CWP) retrievals from the *Geostationary Operational Environmental Satellite-13* (*GOES-13*) imager data are also assimilated during daytime hours (Minnis et al. 2011; Jones et al. 2015). Both radar reflectivity and CWP are assimilated using previously developed forward operators that convert the model state hydrometeor variables into simulated reflectivity and CWP for comparison with observations (e.g., Yussouf et al. 2013; Jones et al. 2015). Assimilation of these data provides the initial conditions of the convective features and the near-storm environment within the model analysis.

Three sets of experiments are conducted for each case. One uses the radiation and aerosol-aware version of the Thompson scheme (THOMP), the second uses the real-time configuration NVD cloud microphysics (NVD-RLT), and the third uses a modified NVD double-moment scheme (NVD-MOD). Details of the modifications are provided in the following section. Otherwise, all experiments use an identical configuration. For all experiments, 3-h forecasts from the first 18 ensemble members are initiated at hourly intervals from 2000 to 0200 UTC. Additional 90-min forecasts are initiated on the half hour beginning at 2030 UTC and ending at 0230 UTC. Forecasts from only the first 18 members are used to reduce computing overhead. Testing has shown that probability statistics do not differ significantly when calculated from 18- or 36-member forecast sets. This forecast configuration mimics the one used during real-time testing in 2017 and 2018.

b. Cloud microphysics schemes

Analysis and prediction of convection within convection-permitting NWP models generally relies on the inclusion of bulk cloud microphysics schemes to represent clouds and precipitation. These schemes characterize hydrometeors in terms of integrated moments of an assumed (usually gamma) size distribution function. Single-moment schemes typically calculate the total mass (mixing ratio), whereas two-moment schemes usually add the total

number concentration, such that the mean hydrometeor diameter is now predicted. Single-moment schemes generally diagnose total concentration (and mean size) by assuming a constant or diagnosed (e.g., temperature dependent) value of the distribution intercept parameter.

The Thompson scheme contains elements of both single- and double-moment configurations as it predicts hydrometeor mass mixing ratios for cloud water, rain, cloud ice, graupel, and snow, but number concentrations for only cloud water, cloud rain, and cloud ice (Thompson et al. 2004; 2008). The Thompson microphysics used here represents the latest version present in the WRF 3.8.1 code distribution, which is radiation aware [i.e., calculates effective radii; Rapid Radiative Transfer Model for GCMs (RRTMG) scheme only] and aerosol aware using monthly climatologies of aerosol concentrations (Thompson and Eidhammer 2014; Thompson et al. 2016). The Thompson scheme employed here also uses the updated number concentration assumptions described by Brown et al. (2016, 2017) to calculate synthetic reflectivity fields. The other scheme used in this research is the NVD scheme, which is a full double-moment scheme and has separate graupel and hail species (Mansell et al. 2010; Ziegler 1985). The NSSL scheme used here is also linked to the RRTMG radiation scheme and uses an initial uniform cloud condensation nuclei (CCN) concentration rather than an aerosol climatology.

Two configurations of the NVD scheme were tested. The first was the same as that used for the 2017 HWT experiment (NVD-RLT). For the second (NVD-MOD), several microphysics parameters were changed to reduce the positive upper-tropospheric cloud bias observed during real-time testing without negatively impacting reflectivity scores. First, the CCN was reduced from 2.0×10^9 to 1.0×10^9 m^{-3}, the latter being the default value (Table 1). Second, the ice hydrometeor fall speed was switched from the default formulation ($49\,420D^{1.415}$; Straka and Mansell 2005) to $42.30D^{0.55}$, which was adapted from Ferrier (1994) and where D is the diameter of the hydrometeor. The lower exponent increases the fall speeds for smaller diameter hydrometeors ($D < 270\,\mu$m), which includes most anvil-level ice crystals. Third, ice and snow hydrometeor fall speeds were increased using scaling factors of 50% and 25% respectively. These changes allow these hydrometeors to fall faster out of the upper-level atmosphere. Finally, maximum graupel and hail-droplet collection efficiencies were boosted from 0.5 and 0.75, respectively, to 0.9, which results in more cloud droplets being scavenged, decreasing the number of frozen droplets in the anvil. These changes act to reduce the total mass and number of snow and ice hydrometeors in the upper troposphere

TABLE 1. Differences in NVD double-moment cloud microphysics parameters between the NVD-RLT and NVD-MOD experiments.

Parameter	NVD-RLT	NVD-MOD
CCN	2.0×10^9	1.0×10^9
Fall speed scheme	Straka and Mansell	Ferrier
Ice fall speed factor	1.0	1.50
Snowfall speed factor	1.0	1.25
Graupel collection efficiency	0.5	0.9
Hail collection efficiency	0.75	0.9

while also increasing their fall speed. The modifications to NVD used in this research were based on several sensitivity studies utilizing idealized deterministic experiments. Further sensitivity studies on multiple real-data cases could not be performed owing to resource constraints and the need to have an improved NVD scheme for cloud analysis in time for summer 2018 real-time experiments. Still, these modifications represent a first step in creating optimal cloud analyses in the NVD scheme, and further optimization tests are likely to occur to prepare for 2019 testing and beyond.

c. Synthetic satellite products

Synthetic satellite brightness temperature and cloud property retrievals are calculated from WRF Model output using the NCEP Unified Post Processing (UPP) system, version 3.1. The UPP contains the Community Radiative Transfer Model (CRTM) that converts the clear- and cloudy-sky-modeled atmospheric state into brightness temperatures for a given satellite sensor and wavelength (Weng 2007; Han et al. 2007). For this research, simulated *GOES-13* infrared ($10.7\,\mu$m) bands are simulated. The UPP code was updated to calculate hydrometeor effective radii required in the cloudy portion of the simulation using formulas specific to each microphysics scheme. Synthetic cloud property retrievals such as cloud-top pressure (CTP) and CWP are calculated directly from the model state using the standard UPP definitions. In the future, it may be possible to apply the retrieval algorithms to the model output to generate fully consistent model and observed retrieval products for comparison, but this task is beyond the scope of this work.

d. Satellite-retrieved products

The cloud properties used for assimilation and verification were derived using the Satellite Cloud and Radiation Property Retrieval System (SatCORPS; https://satcorps.larc.nasa.gov; see Minnis et al. 2008a, 2016) that is based on the algorithms described by Minnis et al. (2011). The assimilated CWP is the value directly computed from the product of the retrieved cloud optical depth τ and hydrometeor effective radius R_e.

Optical depth is retrieved using the Visible Infrared Solar-Infrared Technique (VISST), which computes expected spectral radiances for a range of cloud optical depths and cloud effective radii. The expected spectral radiances are compared with observations and iteratively solved to determine various cloud properties. The cloud properties used for comparison are 10.7-μm brightness temperatures (TB$_{107}$), CWP, and CTP.

In the satellite retrievals, only one phase is assigned to each pixel. Thus, even if an ice-over-water multilayer cloud system or deep convective cloud with a liquid bottom is in the pixel, only one phase (most often ice) is assigned, and τ and R_e are retrieved using models for the selected phase. These are used to compute liquid or ice water path, LWP or IWP, respectively, a value that is assumed to be equal to CWP. For low, all-liquid clouds, the retrieval usually yields a value of LWP, which is within ±30% of microwave instrument retrievals (e.g., Dong et al. 2008; Painemal et al. 2012), even for broken clouds (Painemal et al. 2016). For thin ice clouds, the retrieval tends to slightly overestimate IWP compared to active sensor retrievals (e.g., Mace et al. 1998, 2005). For thick convective clouds, the determination of IWP and CWP from any type of sensor is less certain than for thin clouds. In comparisons with NEXRAD retrievals, Tian et al. (2018) found that the SatCORPS IWP was, on average, 10% lower than the radar estimates for ice-only anvils. However, in stratiform rain portions of the convection, the IWP underestimate is 20% due to the prevalence of larger ice crystals lower in the cloud and to the SatCORPS optical depth retrieval limit of 150. The limit is applied to minimize the false retrieval of very large optical depths at lower sun angles. The SatCORPS IWP retrievals also do not account for the presence of the liquid water in the stratiform portions of the cloud, so the assumption that CWP = IWP in these instances produces an underestimate that will exceed 20%.

To account for this limitation of the satellite retrieval, Smith (2014) developed a parameterization based on combined radar, microwave radiances, and satellite imager data to better estimate CWP. This parameterization first computes the liquid component at the bottom of the thick ice clouds as

$$\text{LWP} = 4.7\tau^{1.1},\qquad(1)$$

and the ice component as

$$\text{IWP}_{\text{iow}} = 0.305\text{IWP}^{1.194}.\qquad(2)$$

The cloud water path then is

$$\text{CWP} = \text{LWP} + \text{IWP}_{\text{iow}}.\qquad(3)$$

For comparison with the models, CWP is computed with Eq. (3) for any pixel having an ice phase, TB$_{107}$ < 253 K, and τ > 10.

The comparisons also employ the satellite-based CTP. For the satellite, it is estimated by first matching the cloud-top effective radiating temperature, which is close to TB$_{107}$ for optically thick clouds, to the lowest altitude having that temperature in a sounding. This cloud effective altitude is assumed to be the cloud-top altitude for water clouds. For ice clouds, the cloud-top altitude is computed using the approaches of Minnis et al. (2008b) and Minnis et al. (2011) for optically thick and thin clouds, respectively. The sounding is constructed by replacing the temperatures in the lower levels (pressure > 700 hPa) of a vertical profile from a numerical weather prediction model with temperatures computed using a regionally dependent apparent lapse rate as described by Sun-Mack et al. (2014). The vertical profiles used here are 6-hourly analyses from the Goddard Earth Observing System, version 5 (GEOS-5), model (Rienecker et al. 2008) interpolated to the time of interest.

3. Object definitions

a. Basic description

Object-based techniques (e.g., Davis et al. 2006a,b; Ebert and Gallus 2009) provide a method for verifying specific features within forecasts against corresponding observations. Advantages of this approach to verification include using object matching to avoid double penalties associated with "close" forecasts (Gilleland et al. 2009), an ability to use forecast and verification fields derived from different data sources, and generation of extensive diagnostic information that allows specific errors to be quantified. A primary limitation of object-based verification is that the highly configurable nature of object identification and matching requires careful selection of subjective thresholds in order to isolate features of interest in the forecast and verification data (Wolff et al. 2014). Object-based verification is particularly attractive for verification of NEWS-e forecasts as it provides a method for verifying imperfect surrogates for convective storm hazards, such as simulated satellite (e.g., Griffin et al. 2017a,b) or radar (e.g., Pinto et al. 2015; Cai and Dumais 2015; Skinner et al. 2018) fields, against corresponding observations.

b. Observed and model object definitions

1) RADAR REFLECTIVITY OBJECTS

The composite reflectivity fields in NEWS-e and MRMS are used as forecast and verification datasets,

respectively, to assess the skill of NEWS-e convective forecasts. Synthetic reflectivity forecast objects are identified using a fixed threshold of 45 dBZ for each microphysics configuration considered. However, an "apples to apples" comparison to MRMS composite reflectivity is not possible owing to differences in the calculation of synthetic reflectivity within varying microphysical parameterizations as well as differences in sampling and interpolation within the MRMS products. Therefore, model climatologies (e.g., Sobash et al. 2016) of NEWS-e forecasts and corresponding MRMS observations are constructed for all available real-time cases in 2017. These climatologies allow extreme percentiles to be matched between MRMS observations and NEWS-e forecasts and MRMS thresholds corresponding to 45 dBZ in NEWS-e forecasts to be identified (Skinner et al. 2018). MRMS thresholds used in this study are 39.73 (41.24) dBZ for cases using Thompson (NVD) microphysics, respectively. The final NEWS-e and MRMS reflectivity object fields are generated by applying the appropriate threshold to unsmoothed values, merging objects with a minimum (boundary) displacement less than 12 km, and applying a size threshold of 144 km^2.

Once composite reflectivity objects have been identified in NEWS-e and MRMS data, they are matched using a total interest score (Davis et al. 2006a) that weights the spatiotemporal displacement between forecast and observed objects. Specifically, the average of ratios of the centroid and minimum displacement between objects to a maximum allowable displacement of 40 km is multiplied by a temporal displacement ratio with a maximum offset of 25 min to calculate the total interest score. Objects with a total interest greater than 0.2 are identified as matches, which corresponds to maximum allowable spatial (temporal) offsets of 32 km (20 min). A complete description of the object identification and matching methodology is provided in Skinner et al. (2018).

The highly configurable nature of object identification and matching is a limitation of object-based verification. It is noted that object-based verification scores are sensitive to varying threshold choices; however, qualitative comparisons between different cases or system configurations remain similar [see the appendix in Skinner et al. (2018)].

2) SATELLITE OBJECTS

Three satellite object types are defined for this research and include TB$_{107}$ < 225 K, CTP < 225 hPa, and CWP > 1.0 kg m^{-2}. The first two object types generally represent the locations of cirrus clouds produced by convection. The CWP type represents this along with other nonconvective clouds in the domain. The CTP and TB$_{107}$ thresholds were defined by analyzing the frequency histograms and determining the values below which the peak occurrence of convective clouds exists. This value varies from case to case and may not capture all objects associated with low-top convection, such as those present in the 17 May case. Using thresholds tailored to a specific case will improve its individual skill scores, but would complicate the comparison of aggregate skill scores computed over all cases. For CWP, the threshold was set to 1.0 kg m^{-2}, which represents the approximate value separating the lowest 90% of CWP values from the highest 10% so that thick clouds and convection-related CWP objects are emphasized. Varying these thresholds by ±10% did not significantly impact the overall results. Future research will refine the satellite object thresholds as more cases and more variables become available (e.g., tropopause height could be used to define an adaptive CTP threshold).

Owing to the different spatiotemporal scales of radar reflectivity and satellite objects, different matching algorithms are used. Satellite objects are matched if the minimum distance between observed and synthetic objects is less than the radius of the major object axis for an observed object with a maximum radius threshold of 400 km. The minimum displacement for satellite objects is 25 km. A fixed radius is not practical for cloud objects since their length radii can range from very small (~10 km) to very large (>500 km) over the life cycle of a single convective storm. Model objects are generated from each member at 15-min intervals over the duration of the forecast period for each event. For each matched object, the object area (km^2), maximum intensity, major and minor length radii (km) are calculated. As with reflectivity objects, the average of ratios of the centroid and minimum displacement between objects is multiplied by a temporal displacement ratio with a maximum offset of 25 min to calculate the total interest score and those with a total interest greater than 0.2 are identified as matches. The larger the analyzed or observed cloud, the larger the object area should be. This matching methodology allows matched pairs (in both reflectivity and satellite fields) to have different sizes and intensities. The simple total interest function used for matching only considers the centroid and minimum (boundary) displacement in space and time.

Synthetic satellite objects are generated directly from the postprocessed 3-km-resolution model output. Prior to the creation of the observed satellite objects, *GOES-13* TB$_{107}$ and retrievals are objectively analyzed onto a 5-km grid and the appropriate parallax corrections are applied to cloudy pixels. An example of TB$_{107}$, CWP, and CWP objects is provided in Fig. 1, which shows the *GOES-13*
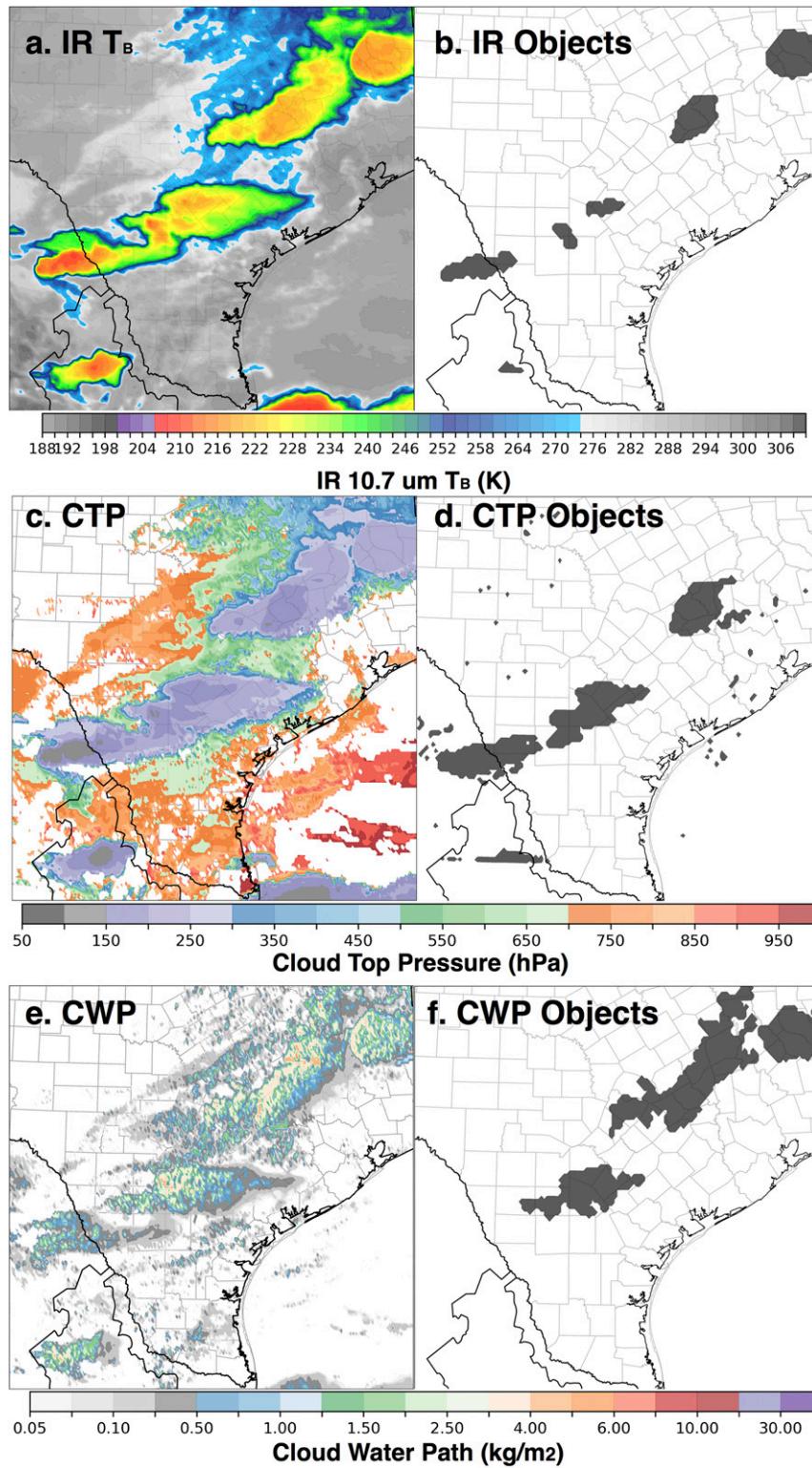
FIG. 1. *GOES-13* (a) infrared (10.7 μm) imagery, (c) CTP, and (e) CWP at 2300 UTC 23 May with corresponding objects shown as gray regions in (b), (d), and (f).
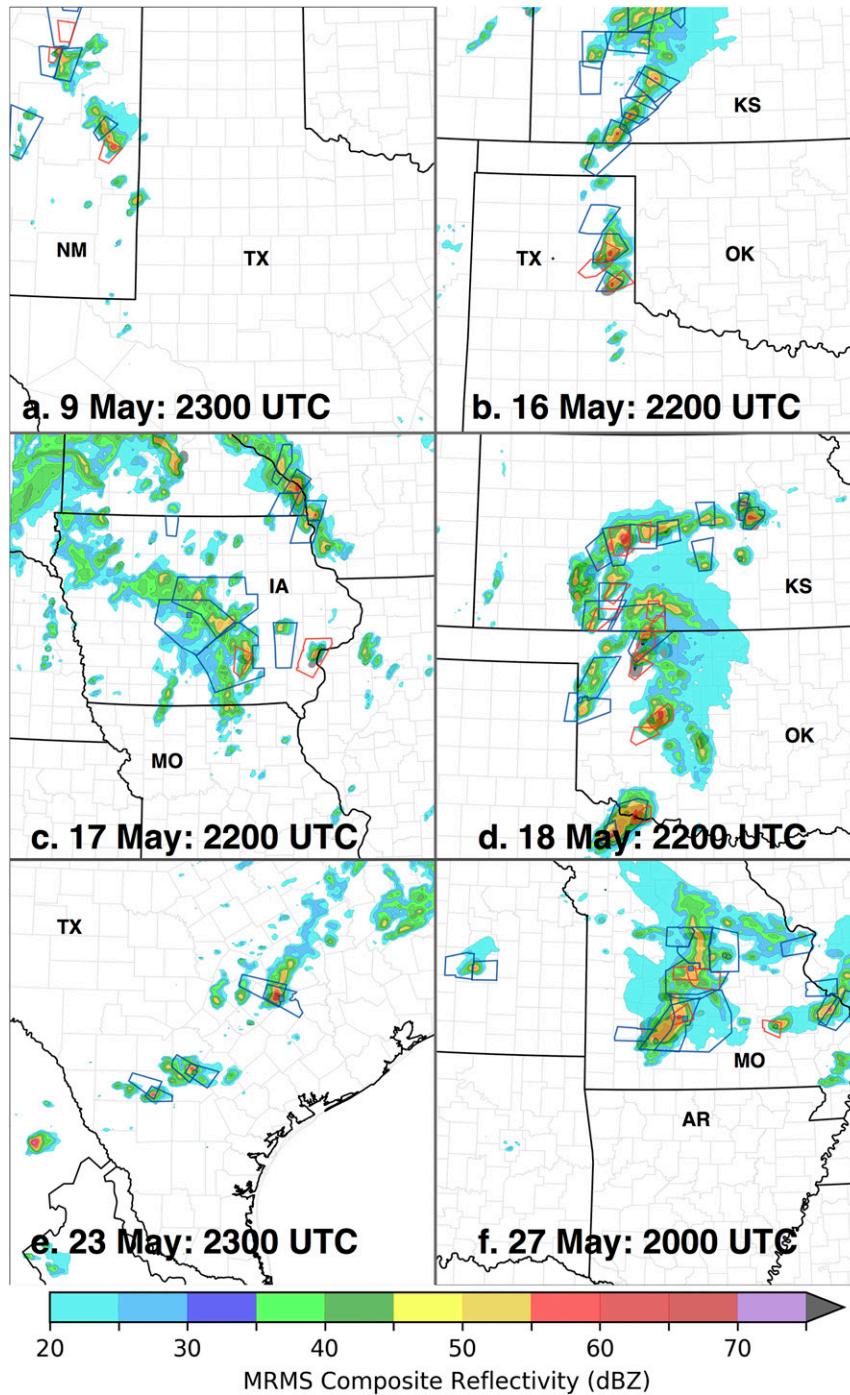
FIG. 2. (a)–(f) MRMS composite reflectivity at a selected analysis time for each case showing severe thunderstorm (blue) and tornado (red) warnings valid at these times. Local time (central daylight time) is UTC − 5 h.

imager data at 2300 UTC 23 May for each field along with the corresponding objects calculated using the method described above. At this time, several areas of developing convection are present in southeastern Texas (TX),

defined by several areas of low TB_{107} and high CTP. The TB_{107} and CTP objects correspond well in both size and orientation to these areas (Figs. 1a–d). For CWP, the convective cirrus and lower-level clouds associated with

developing convection act to create one large object where two or three exist for TB_{107} and CTP (Figs. 1e,f). In other cases, CWP objects associated with nonprecipitating midlevel stratus and cumulus clouds not associated with convection are generated (not shown).

4. Case overviews

a. Descriptions

Each May 2017 case generated numerous severe weather warnings and corresponding reports, though each case had unique characteristics. Figure 2 shows MRMS composite reflectivity at a selected time for each case, with corresponding valid severe weather warnings overlaid. The coverage of strong convection varies from cases to case with the greatest number of isolated supercells occurring on 16 and 18 May (Figs. 2b,d). These two cases also generate the most tornado and severe hail reports of the six being studied (Table 2). Fewer isolated severe storms exist for 9 and 23 May and the number of severe weather reports associated with these events remains small. The other two cases, 17 and 27 May, generate a more linear mode of convection, compared to the isolated characteristics of the other cases, and also generate large numbers of severe wind reports (Figs. 2c,e; Table 2). For 17 May, the maximum measured reflectivity values are lower than the other cases with the overall depth of the convection being lower (not shown). Despite this, more severe weather reports are generated on this day than any of the others.

Corresponding *GOES-13* TB_{107} imagery at the same time for each case highlights additional case-specific characteristics (Fig. 3). Cold cloud tops ($TB_{107} < 230$ K) are associated with the severe convection for the 18 and 27 May cases and generate the coldest and largest cirrus coverage, with areas of $TB_{107} < 210$ K present (Figs. 3d,f). The cirrus coverage for 9 and 16 May is smaller and corresponds to an early phase of storm development at these times, which expands in coverage in later hours (not shown). Finally, the 23 May and particularly the 17 May cases generate warmer cloud tops than the other cases. This indicates shallower convection, but the number of severe weather reports indicates that this convection nevertheless produced extensive severe weather.

b. Histogram analysis

Before objects can be defined, it is important to understand the distribution of observed and simulated satellite data for each event. To visualize these distributions, frequency histograms of observed and simulated 10.7-μm brightness temperatures (TB_{107}), CTP,

TABLE 2. Numbers of tornado, severe hail (diameter > 1.0 in.), and high-wind reports associated for each case between 1800 and 0500 UTC the following day within the model domain.

Date	Tornadoes	Hail	Wind
9 May	2	12	1
16 May	20	67	3
17 May	11	40	160
18 May	15	86	29
23 May	0	12	8
27 May	8	71	92

and CWP were created for each case. These histograms represent an aggregate of observations and ensemble mean analyses at 30-min intervals starting at 2000 UTC and ending at 0230 UTC each day. Data are binned into 2-K intervals and normalized to the percent of observations or grid points present within a particular bin. For all cases, the TB_{107} histogram indicates a bimodal distribution with one grouping corresponding to upper-level ice-phase clouds ($TB_{107} < 230$ K) and one corresponding to mostly clear conditions ($TB_{107} > 275$ K) (Fig. 4). The exact distribution of both observed and simulated TB_{107} varies from case to case and differences between each experiment are also apparent. On 9 May, the majority of the domain is relatively clear with most observed and simulated TB_{107} results being greater than 275 K. The secondary peak ($TB_{107} < 225$ K) corresponds to the cirrus clouds generated from convection, which have a much smaller areal coverage (Fig. 3a). Overall, the observations and simulated model data agree, but one important difference is present. Note that the NVD-RLT experiment generates more and colder TB_{107} values compared to observations or the other experiments. Similar results were present for 16 May except that the domain was more evenly balanced between convection and clear-sky areas (Fig. 4b). The largest differences between the observations and model output occurs on 17 May, with all experiments greatly overestimating the distribution of cold ($TB_{107} < 230$ K) cloud tops relative to the observations while underestimating clear-sky areas (Fig. 4c). The NVD-RLT experiment is by far the worst performer, with both THOMP and NVD-MOD generating somewhat warmer clouds and over a larger temperature range. The 18 May case is dominated by very large convective cirrus which covers over 80% of the domain by 2300 UTC (Figs. 3c and 4d). Note that TB_{107} is colder than the previous cases, indicating deeper convection and higher cirrus and anvil clouds. As before, NVD-RLT has a cold bias compared to observations and the other experiments. The 23 May case shares many of the characteristics of 9 May except for a drop in TB_{107} between 285 and 295 K in simulated data
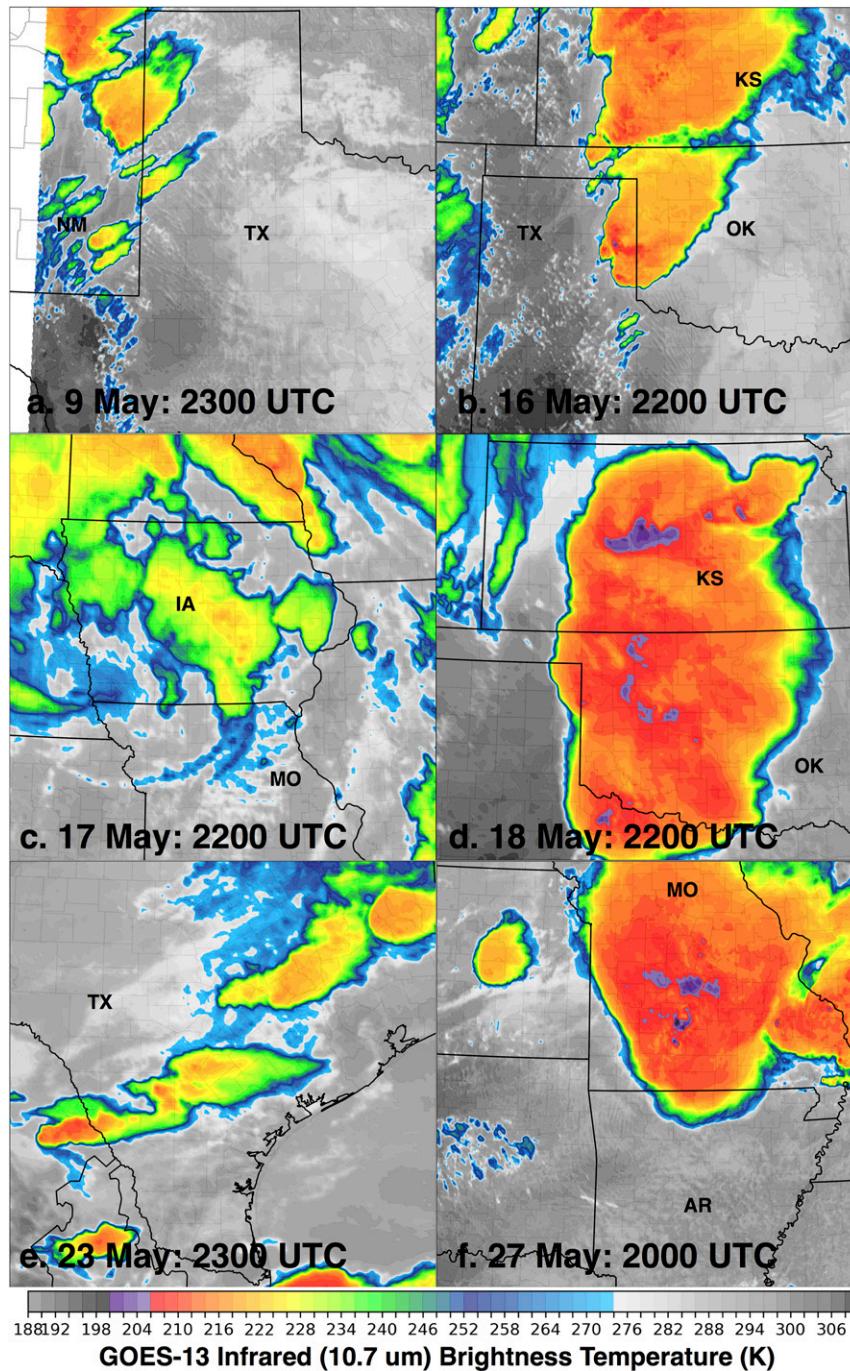
FIG. 3. (a)–(f) *GOES-13* infrared (10.7 μm) imagery for each case corresponding to the times in Fig. 2. Local time (central daylight time) is UTC − 5 h.

that are not present in the observations (Fig. 4e). This is attributable to a bias in the CRTM-derived TB_{107} over the Gulf of Mexico. The 27 May case is similar to the 16 May case except that cloudy TB_{107} is somewhat colder (Fig. 4f). Overall, it is clear that NVD-RLT generates simulated deep-convective TB_{107} values that

are both too cold and present over a larger area compared to the observations and other experiments.

Similar patterns are present in the CTP histograms with NVD-RLT generating higher clouds (lower CTP) for several cases (Fig. 5). CTP is binned into 25-hPa intervals and clear-sky areas (CTP = null) are not
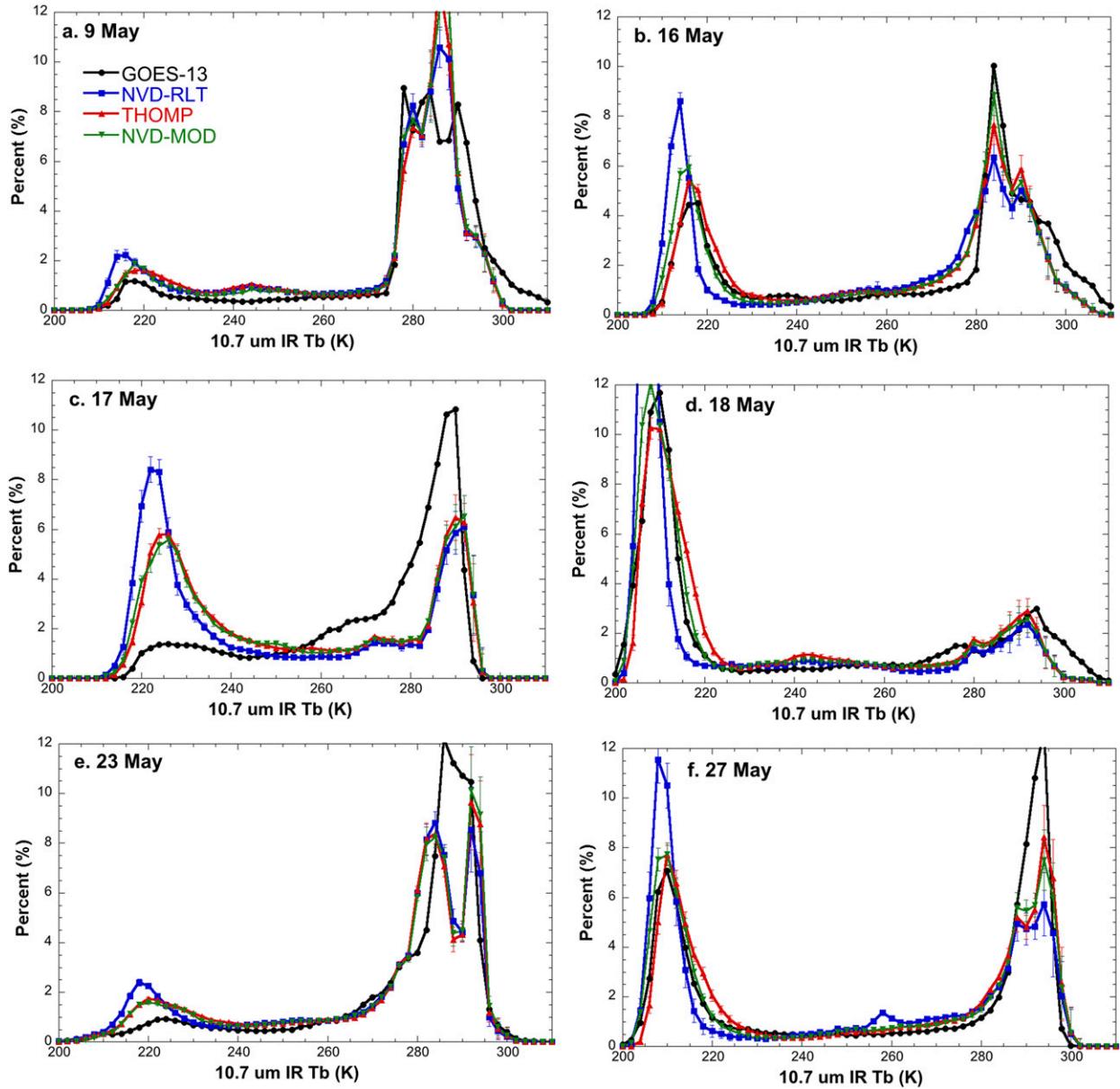
FIG. 4. (a)–(f) Frequency histogram of *GOES-13* and ensemble mean simulated TB_{107} aggregated over analysis times between 2000 and 0230 UTC at 30-min intervals for each case. Colder TB_{107} values indicate cirrus outflow from convection, and warm TB_{107} indicates clear-sky areas. Vertical lines indicate the standard deviation of TB_{107} over all ensemble members for each bin.

included in the normalized percentages. Model differences are particularly evident for 16, 17, 18, and 27 May. Experiments that generate higher cloud tops are also those that generate the colder TB_{107}. For higher clouds below the tropopause, the surrounding environment would be colder, corresponding to the colder clouds observed and simulated by the model. In the case of nonconvective clouds with lower cloud tops, some differences exist between the retrievals and simulated CTP distributions. For the 9 and 23 May cases, the simulated

nonconvective clouds have a peak coverage roughly 50 hPa below the observed values. Some of the bias may be due to observational error since the differences are within the range of the observational and model uncertainties. Changes to the cloud microphysics did not impact the height distribution for these low-level clouds, but in the case of NVD-MOD, they did increase their coverage for 9 and 16 May (Figs. 5a,b).

Finally, CWP histograms in Fig. 6 differ significantly from TB_{107} and CTP due to characteristics specific to the
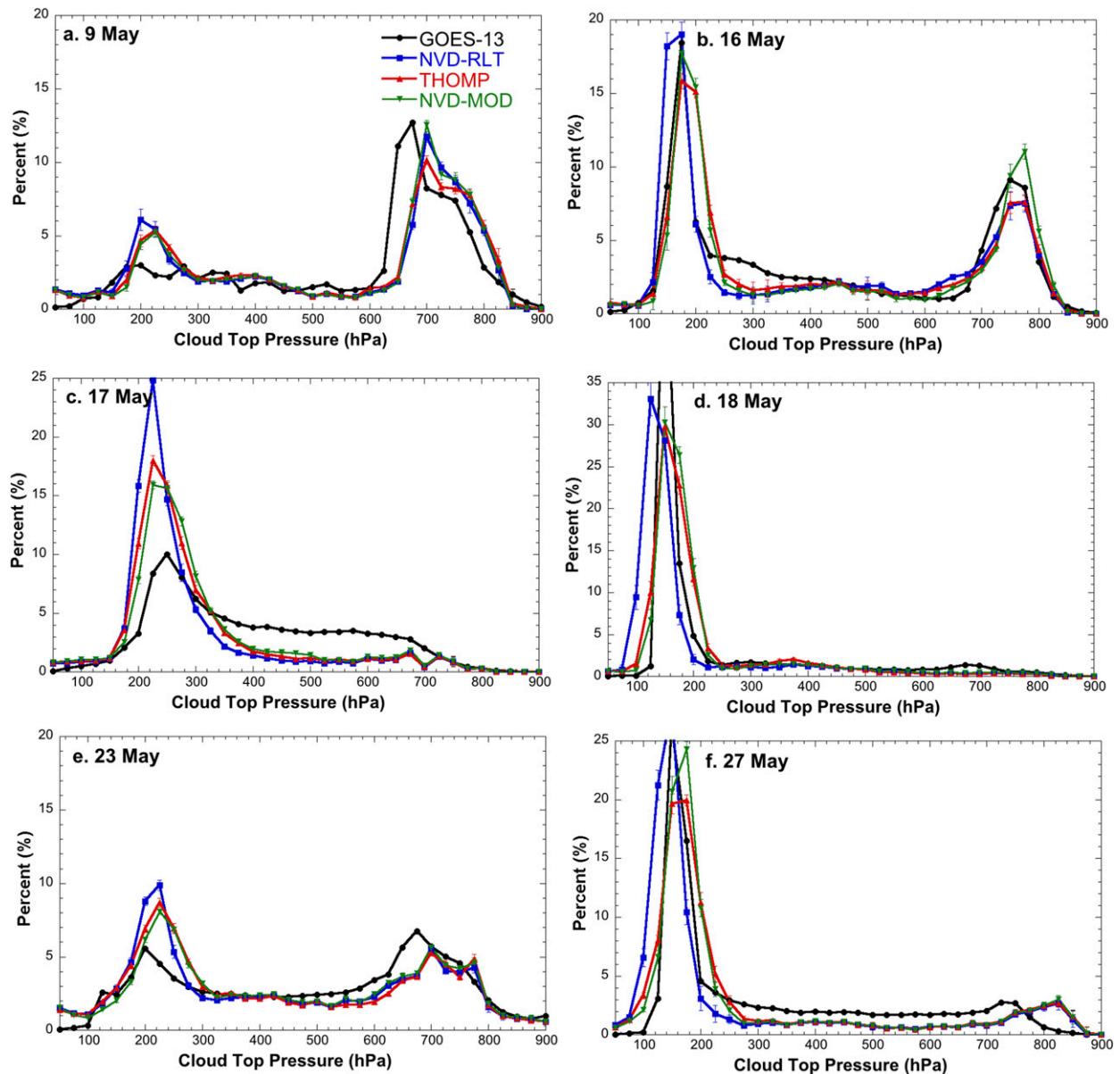
FIG. 5. As in Fig. 4, but for CTP (hPa).

calculation of the cloud water path as noted earlier. Specifically, the value of CWP is generally not sensitive to height, so the microphysics dependencies observed for CTP and TB_{107} are often not present. The occurrence of CWP is greatest for very small values ($<0.5 \, kg \, m^{-2}$) and rapidly decreases thereafter. All experiments generate similar distributions for CWP $> 2.0 \, kg \, m^{-2}$ while differences for larger CWP values are very small ($<0.1\%$). However, the observed CWP distribution differs significantly from the model analyses in several cases. For all cases except 18 May, the coverage of CWP $< 2 \, kg \, m^{-2}$ is greater in the model output compared to the observations.

The 16, 18, and 27 May experiments generate lower occurrences of CWP $> 2.0 \, kg \, m^{-2}$ compared to the retrievals with the caveat that for 18 May the occurrence of retrieved CWP reaches zero near CWP $= 5 \, kg \, m^{-2}$. This is not unexpected since the CWP algorithm retrieval is constrained to $\tau \leq 150$, corresponding to moderate-to-heavy precipitation, so that, even with the parameterized correction of LWP, it peaks between 3 and $5 \, kg \, m^{-2}$ depending on the atmospheric conditions and the retrieved value of R_e. The maximum values vary from 4 to $6 \, kg \, m^{-2}$. The model-simulated CWP does not have this limitation
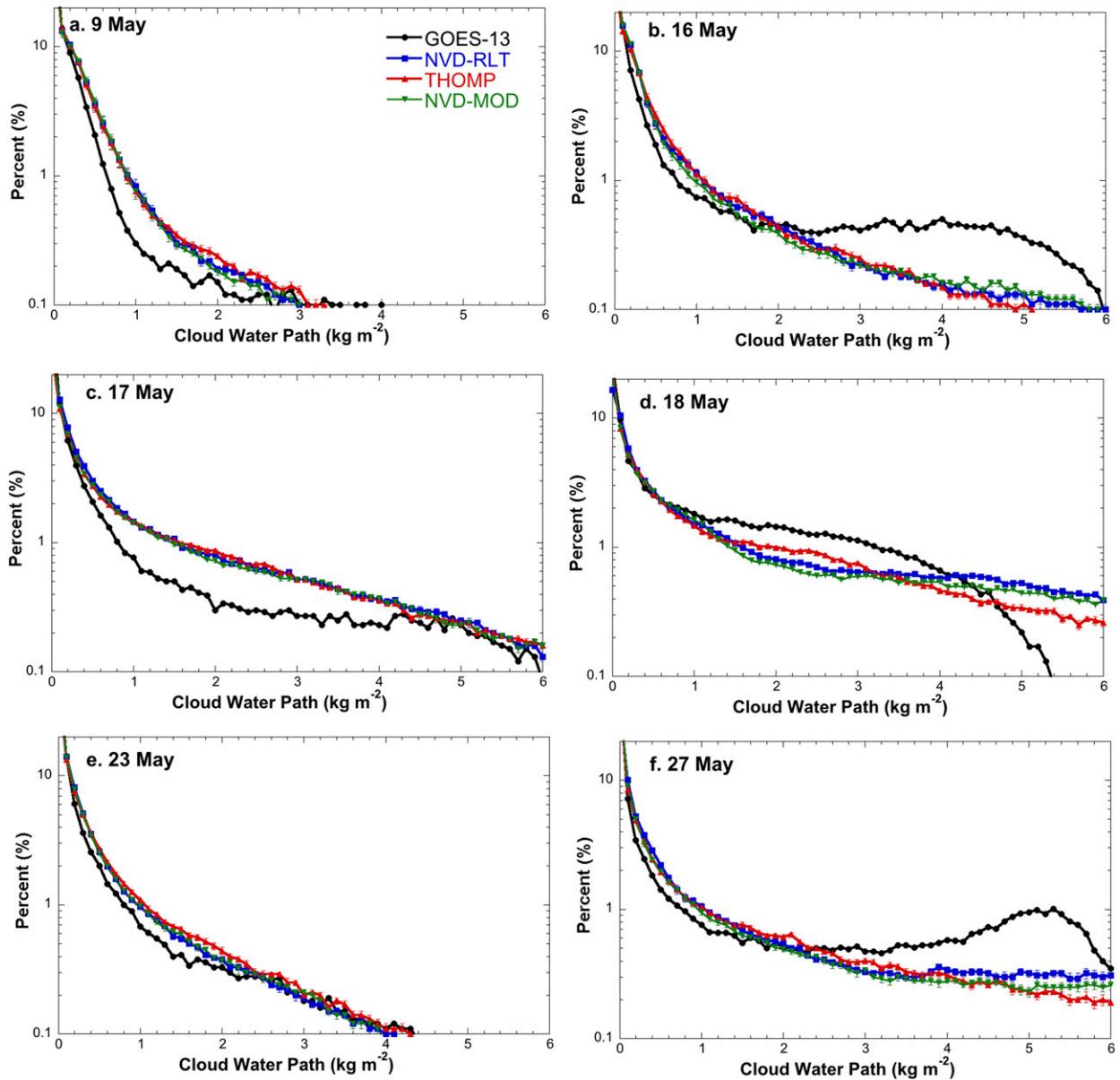
FIG. 6. As in Fig. 4, but for CWP (kg m^{-2}). Note that the y axis is on a logarithmic scale to highlight small differences in model distributions for CWP > $1.0\,\text{kg m}^{-2}$.

and can generate values in excess of $20\,\text{kg m}^{-2}$ in small areas near convective cores.

5. Object verification

a. Example forecasts

To visualize the impact of applying different cloud microphysics schemes to modeled cloud properties, two analysis examples are provided. At 2300 UTC 9 May, simulated ensemble mean TB_{107} indicates several existing and developing storms in eastern New Mexico

(NM) and western TX (Fig. 7). Compared to the observed TB_{107} at this time (Fig. 2a), all experiments overestimate the cirrus cloud coverage, with NVD-RLT generating the largest overestimate. For example, NVD-RLT generates a large area of $TB_{107} < 225\,\text{K}$ in northwest TX associated with the supercell in northeastern NM, while the eastern extent of the cirrus in THOMP and NVD-MOD is significantly less (Figs. 7a–c). Calculating the difference (THOMP − NVD-RLT and NVD-MOD − NVD-RLT) in TB_{107} shows that both THOMP and NVD-MOD generate much higher (warmer) TB_{107}
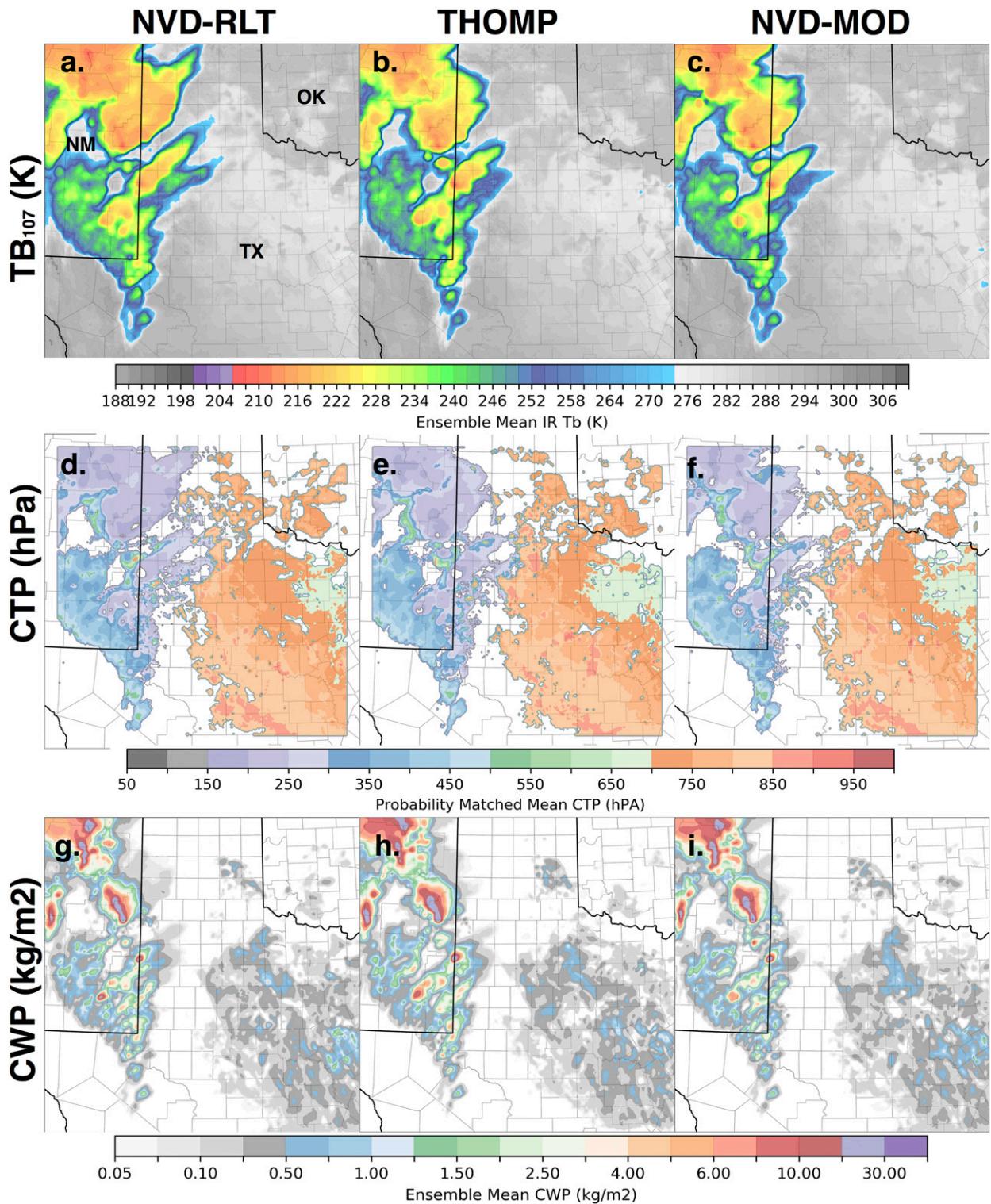
FIG. 7. Ensemble mean simulated (a)–(c) TB_{107}, (d)–(f) CTP, and (g)–(i) CWP at 2300 UTC 9 May for the (left) NVD-RLT, (center) THOMP, and (right) NVD-MOD experiments. Note the greater coverage of $TB_{107} < 220$ K and CTP < 225 hPa in the NVD-RLT experiment compared to the others.
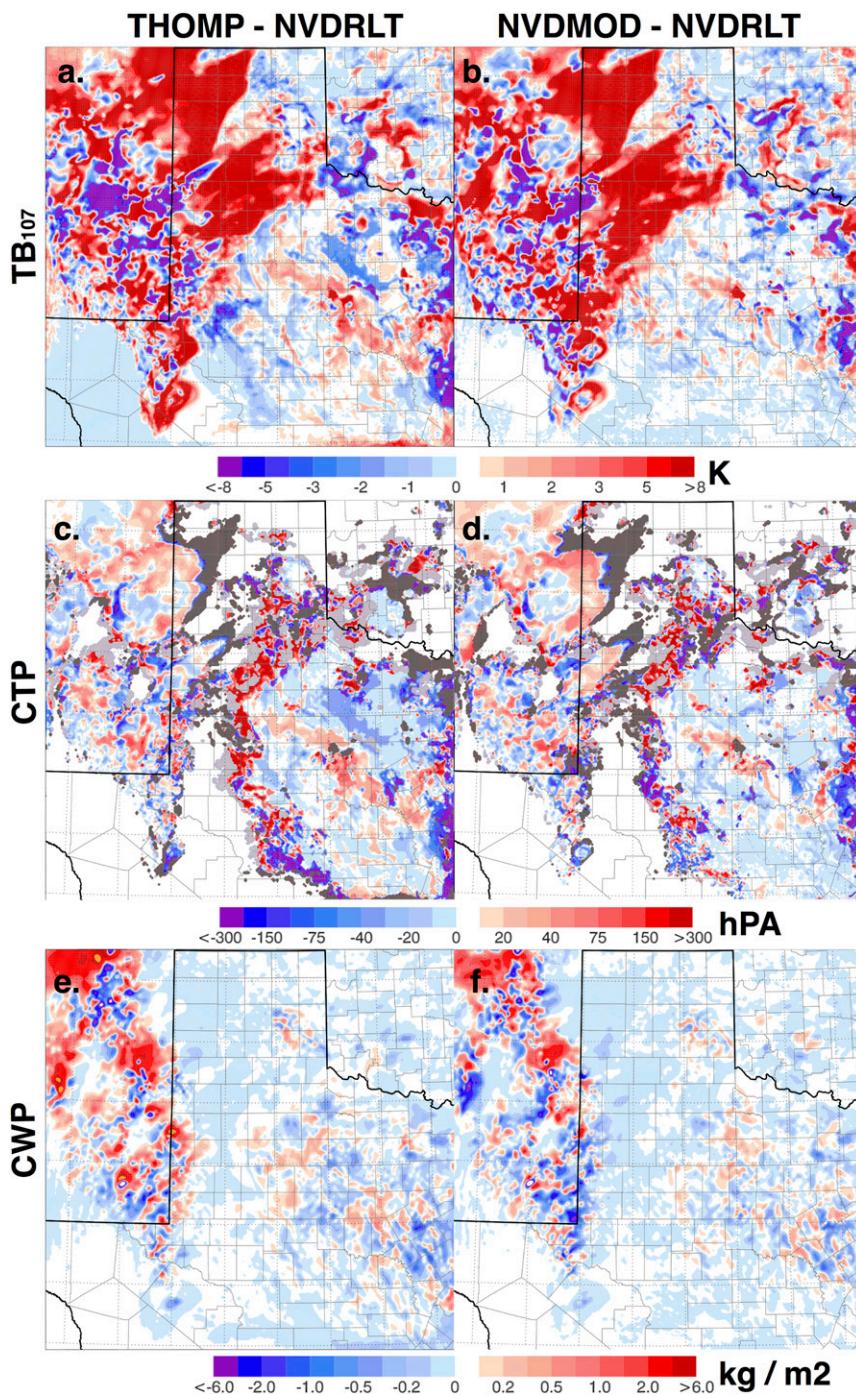
FIG. 8. Differences in (a),(b) TB_{107}, (c),(d) CTP, and (e),(f) CWP between (left) NVD-RLT and THOMP and (right) NVD-RLT and NVD-MOD at 2300 UTC 9 May. For TB_{107}, red colors indicate that either THOMP or NVD-MOD generates warmer TB_{107} than NVD-RLT, indicating less cloud cover. Blue colors indicate the opposite. CTP is similar except that gray colors indicate where NVD-RLT generates CTP retrievals where none are present in either THOMP or NVD-MOD.
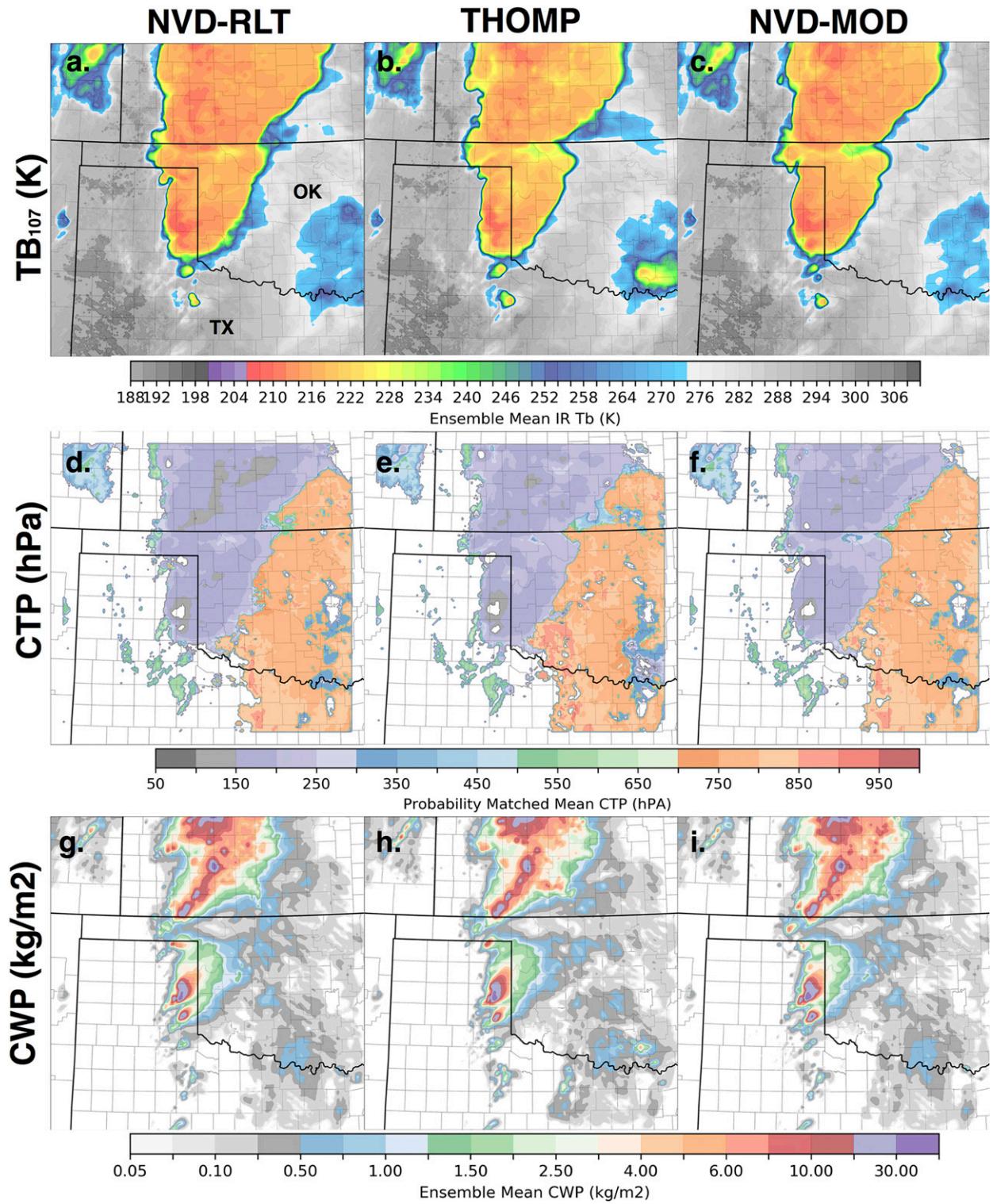
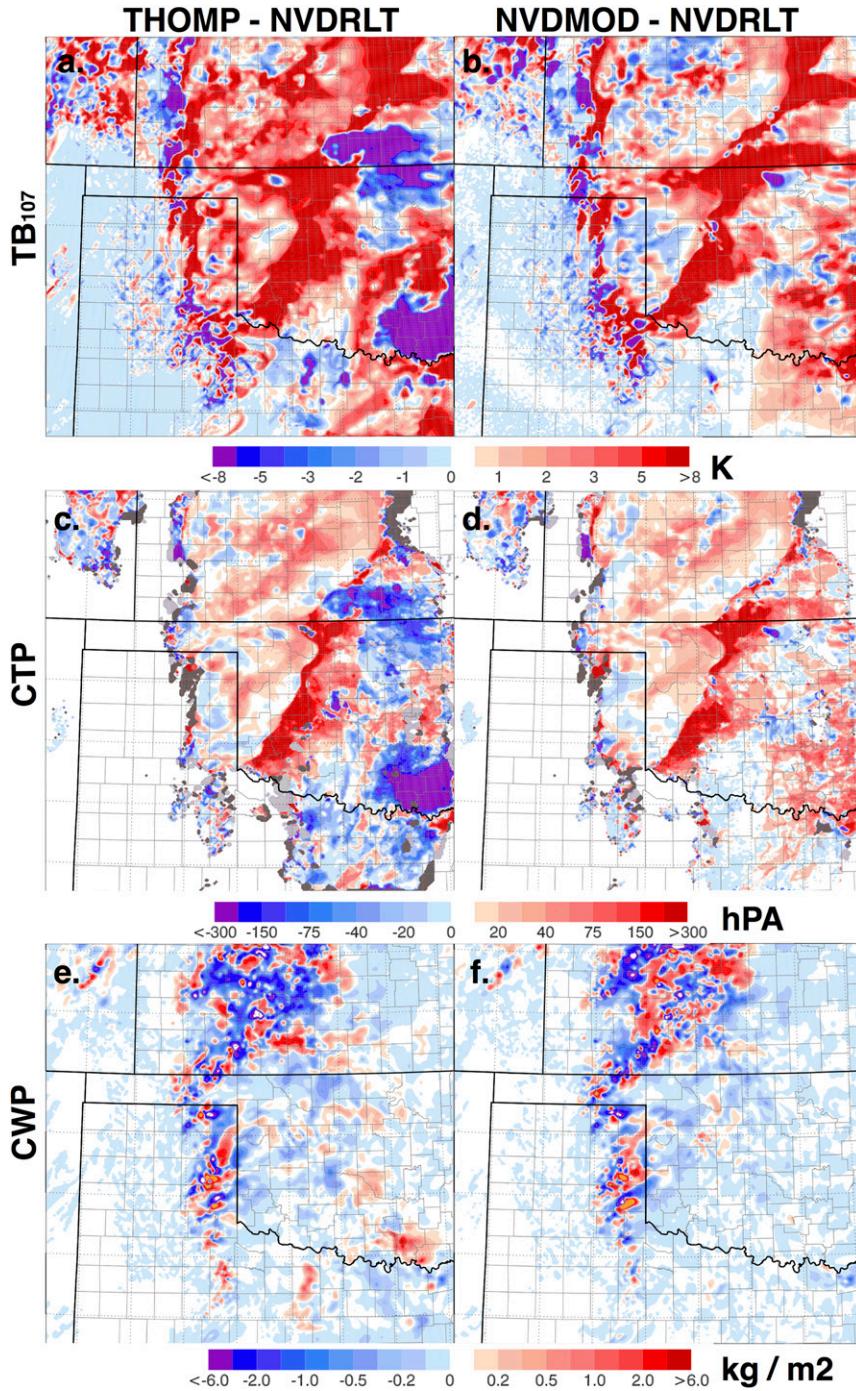FIG. 9. As in Fig. 7, but at 2200 UTC 16 May.

FIG. 10. As in Fig. 8, but at 2200 UTC 16 May.

values than NVD-RLT ahead of the ongoing convection indicating the cirrus shield generated by THOMP and NVD-MOD is much smaller (Figs. 8a,b). Other differences exist where all experiments generate cloud cover, but the differences are more random in nature.

Differences in ensemble mean CTP are similar, with NVD-RLT generating the highest cloud tops among all of the experiments at this time (Figs. 7d–f and 8c,d). The overprediction bias of NVD-RLT over THOMP and NVD-MOD is further illustrated by the gray regions in Figs. 8c and 8d, which denote areas where either
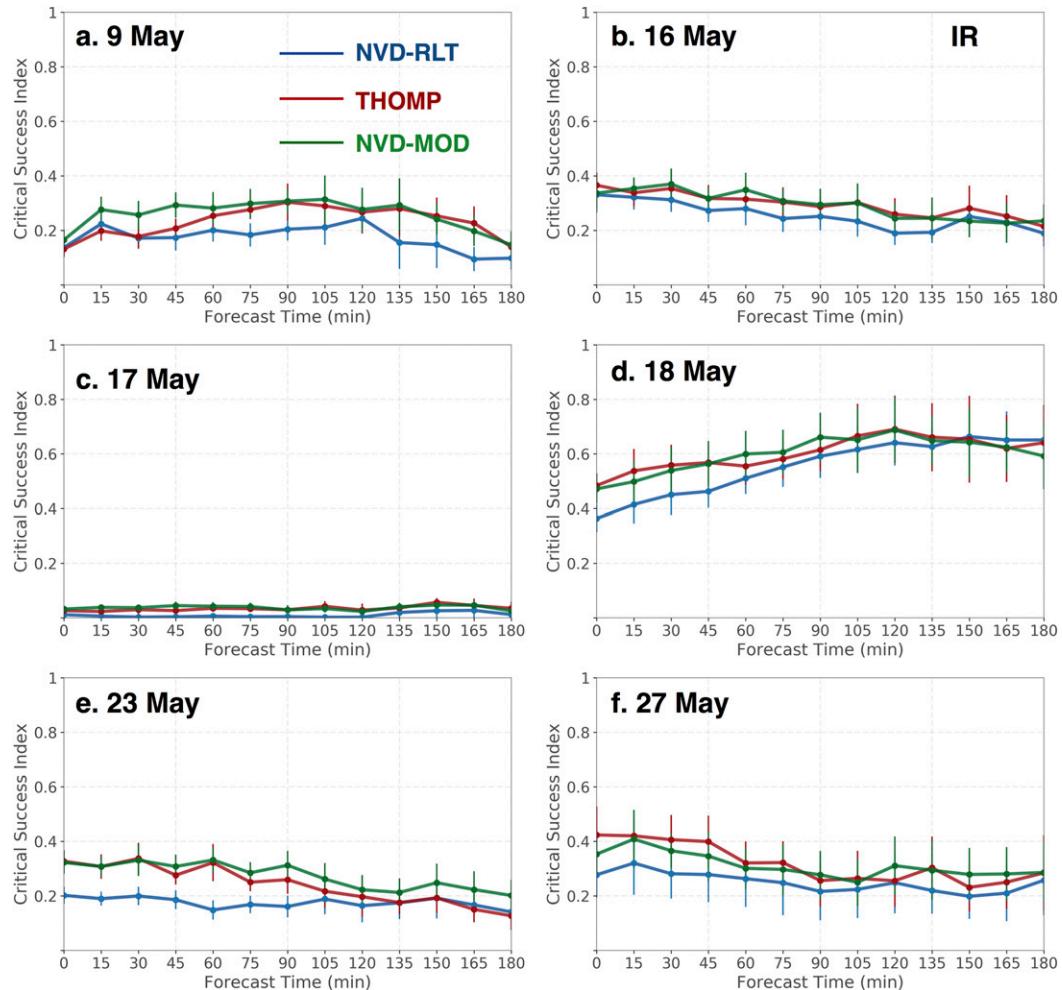
FIG. 11. Ensemble mean CSI as a function of forecast time averaged over all forecasts for each case for TB_{107}. The blue color represents NVD-RLT, red represents THOMP, and green represents NVD-MOD. Vertical lines indicate standard deviations of CSI calculated among all forecast members.

THOMP or NVD-MOD did not generate clouds and NVD-RLT did, precluding the calculation of a difference in CTP. Other differences are present around the edges of the nonconvective cloud field present over most of TX. Inside this region, CTP associated with low-level (CTP < 700 hPa) clouds exhibits much smaller variations between experiments compared to the higher-level clouds associated with convection.

Finally, differences in CWP between each experiment are generally smaller and more isolated in nature. However, the decrease in the eastward extent of CWP < 0.1 kg m^{-2} for THOMP and NVD-MOD is evident (Figs. 9g–i and 10e,f).

Similar impacts from changing cloud microphysics schemes were observed for the 16 May case at 2200 UTC (Figs. 9 and 10). The NVD-RLT experiment generates the coldest TB_{107} values and largest areal coverage of

$TB_{107} < 225$ K compared to either the THOMP or NVD-MOD experiments (Figs. 9a–c). The latter two experiments still show separation in the cirrus clouds from the southern and northern storm complexes present in the observations while it is almost completely gone in NVD-RLT. Difference plots show the large areas of warmer TB_{107} being generated by THOMP and NVD-MOD east of the developing convection compared to NVD-RLT (Figs. 9a–c and 10a,b). The THOMP simulation also generates an area of colder TB_{107} in southeastern Oklahoma (OK) owing to anomalous convection being developed by the model, which is not generated in either NVD experiment. Similar patterns are evident when comparing CTP, with THOMP generating larger differences in the low-level cloud field in eastern OK compared to NVD-MOD (Figs. 9d–f and 10c,d). Finally, differences in CWP were
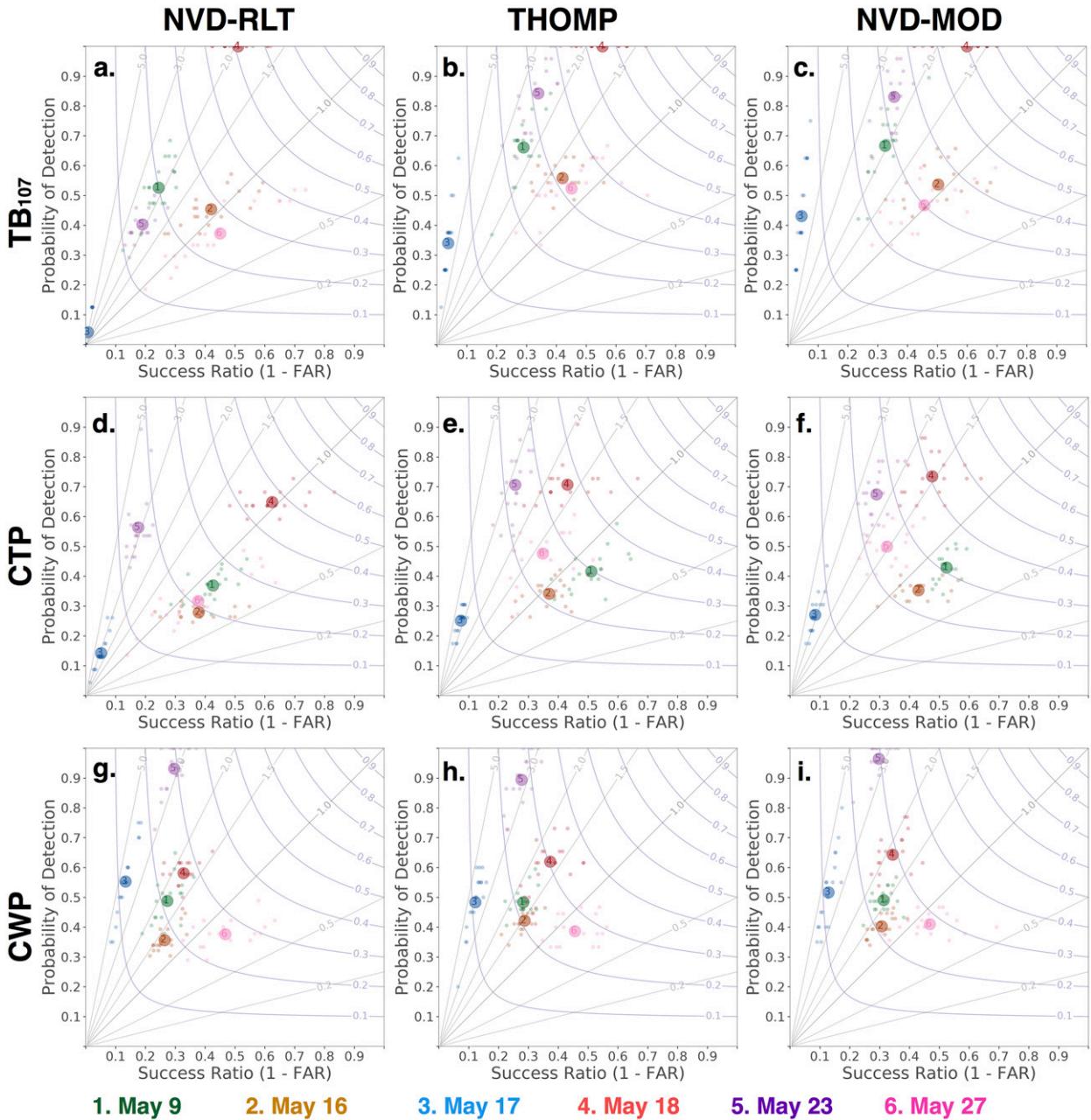
FIG. 12. Performance diagram for (a)–(c) TB$_{107}$, (d)–(f) CTP, and (g)–(i) CWP for each experiment and case for 60-min forecasts. Each case is shown as a different color. Large dots indicate ensemble mean performances while smaller dots indicate individual member performances. In these diagrams, minimum skill is located in the bottom left, with skill maximized in the top right. For a perfect score, POD = 1 and success ratio = 1. Curved blue lines represent CSI, and diagonal gray lines represent bias.

generally smaller overall with larger values confined to areas near convection (Figs. 9g–i and 10e,f).

It is clear from both examples that NVD-RLT overestimates cirrus cloud coverage and cloud height compared to observations while the other two experiments are qualitatively closer to reality. These comparisons are consistent with the frequency histograms shown for

these cases where the distribution of both TB$_{107}$ and CTP from NVD-RLT peaks left of the observed and values from the other experiments.

b. Individual case comparisons

Qualitatively, both THOMP and NVD-MOD generate more-realistic cirrus cloud characteristics than
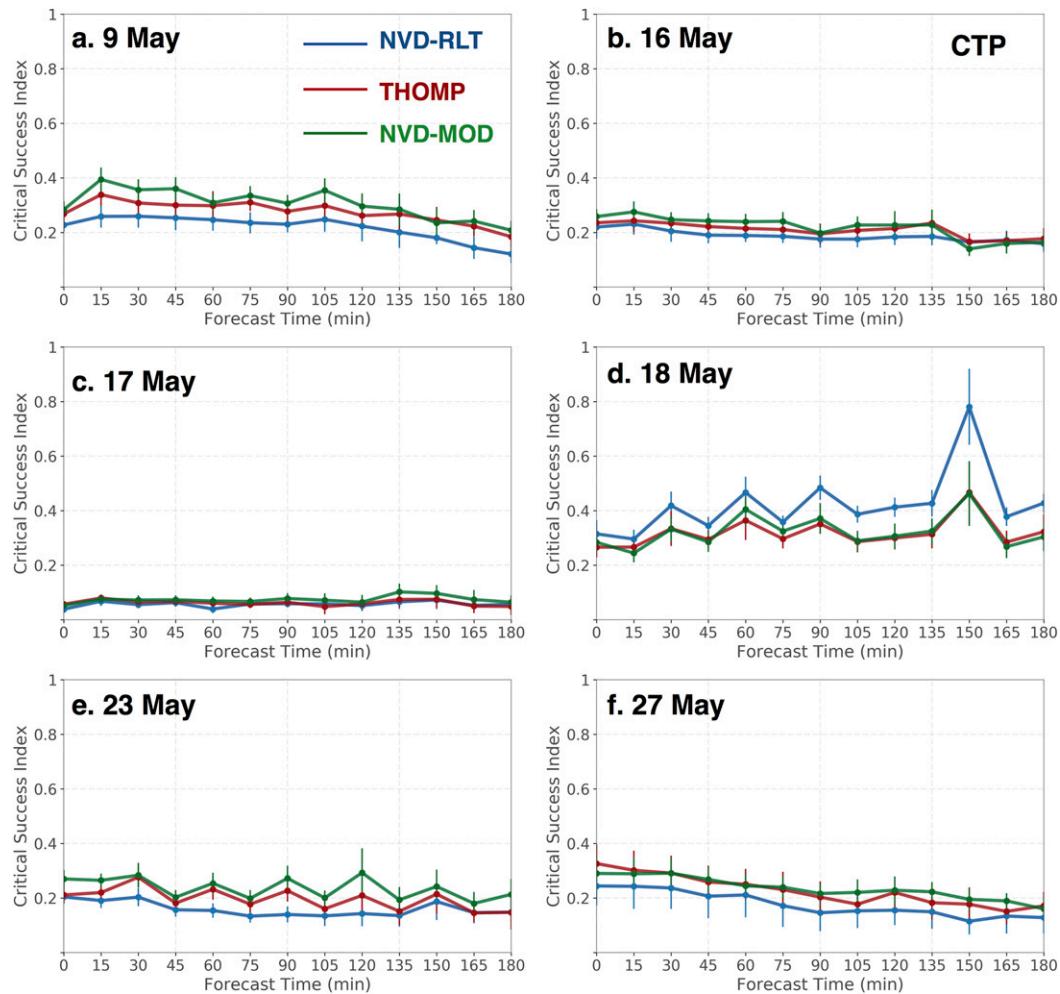
FIG. 13. As in Fig. 11, but for CTP objects.

NVD-RLT, but these differences must be quantified before any final conclusions can be drawn. As a result, object-based skill scores in the form of the critical success index (CSI) were computed for TB_{107}, CTP, and CWP. CSI scores are calculated by aggregating contingency table elements of hits (matched object pairs), false alarms (unmatched NEWS-e objects), and misses (unmatched *GOES-13* objects) across all forecasts for each case. The ratio of hits to the sum of hits, false alarms, and misses is then calculated for each available forecast time (Fig. 11). CSIs vary from experiment to experiment, with 18 May having the highest and 9 and 17 May being the lowest. CSI generally decreases as a function of forecast time in all cases except 18 May. This event was associated with a small number of very large cloud objects at later forecast times, resulting in very high POD values compared to the other experiments (not shown). Overall, CSI is generally related to the convective mode of the individual event with those associated

with intense, isolated supercells performing best. CSI varies substantially as a function of experiment with NVD-RLT generally performing the worst and either THOMP or NVD-MOD performing the best. Some overlap between members of each experiment exists, but most members lie outside the spread envelope of the other experiments at least out to the 2-h (120 min) forecast time. To better visualize the overall performance at a specific forecast time (60 min), performance diagrams (Roebber 2009) relating probability of detection (POD), false alarm ratio (FAR), CSI, and frequency bias are generated (Skinner et al. 2018). More skillful forecasts are located in the top right of the diagram and less skillful forecasts to the bottom left. Figure 12 shows performance diagrams generated for each case and experiment. Significant spread exists between each case with ensemble mean POD ranging from near 1.0 for the 18 May case to as low as 0.1 on 17 May (Fig. 12a). Due to the relatively high FAR values, the

success ratio is generally below 0.5. Skill scores are improved in both THOMP and NVD-MOD, with much of the improvement coming from an increase in POD and somewhat smaller decreases in FAR. All cases show at least some improvement in THOMP and NVD-MOD compared to NVD-RLT, with the largest improvements in the 17 and 23 May cases (Figs. 12b,c).

Similar impacts on forecast skill were observed for CTP (Fig. 13). Overall, CSI values were generally lower than the TB_{107} values, which is largely attributable to an increase in FAR from smaller CTP objects. NVD-MOD outperforms NVD-RLT for all cases except for the 18 May case where the skill is somewhat reduced at all forecast times (Fig. 13c). This event has the fewest CTP objects of all of the cases and is more sensitive to smaller differences in object matches versus false alarms (Table 3). Closer inspection of these cases showed that a few large CTP objects classified as matches in NVD-RLT were classified as false alarms in the other experiments due to the distance threshold being exceeded by a small amount. Corresponding 1-h forecast performance diagrams also show the greatest improvement associated with the 16 and 23 May events (Figs. 12d–f). The very low overall skill for the 17 May case is a result of low POD values resulting from the choice of TB_{107} and CTP thresholds used in object classification. Tuning the thresholds for this case does improve its skill scores, but at the expense of degrading the scores for the other cases. Skill is slightly reduced by THOMP and NVD-MOD compared to NVD-RLT for the 18 May event, as noted above.

Finally, the results from comparing CSI for CWP are somewhat different than those for TB_{107} and CTP. For most cases, the differences in CSI are small when comparing each experiment, and do not change significantly as a function of time (Fig. 14). CSIs from ensemble members associated with each experiment often overlap with those from the other experiments. Performance diagrams for 1-h forecasts show similar results (Figs. 12g–i), with only small differences between different cloud microphysics schemes, which is consistent with the qualitative examples shown above (Figs. 8 and 10).

c. Combined case summary

Skill scores were aggregated across the entire sample of cases to assess bulk differences in forecast skill. Figure 15 shows performance diagrams for IR, CTP, and CWP objects at $t = 60$, 120, and 180 min. Both THOMP and NVD-MOD increase in CSI and POD with decreasing FAR compared to NVD-RLT for the first 2 h of the forecast period. Bias does increase in THOMP and NVD-MOD, corresponding to an increase in the total number of forecast objects from 39 674 in NVD-RLT to 47 054 and 46 692 in THOMP and NVD-MOD, respectively, at $t = 60$ min (Table 3). At this forecast time,

TABLE 3. Observed and simulated forecast satellite objects for each case. The number of model objects represents a total from all ensemble members.

Event	IR	CTP	CWP	REFL
9 May				
Observations	209	637	365	1535
NVD-RLT	7120	9016	11 890	27 700
THOMP	7574	8966	10 777	34 270
NVD-MOD	6958	8916	10 513	30 839
16 May				
Observations	387	623	607	2673
NVD-RLT	6985	7897	12 644	35 924
THOMP	8175	9606	13 544	41 058
NVD-MOD	7163	9429	12 402	44 262
17 May				
Observations	110	314	261	2590
NVD-RLT	7643	11 635	14 224	54 413
THOMP	11 977	13 033	13 319	69 183
NVD-MOD	13 257	12 682	13 559	57 598
18 May				
Observations	124	293	352	4571
NVD-RLT	5002	4353	8402	91 244
THOMP	4441	6835	7853	105 297
NVD-MOD	4468	7097	9014	103 287
23 May				
Observations	330	436	283	1479
NVD-RLT	8755	15 101	12 087	31 486
THOMP	10 053	12 669	11 013	36 998
NVD-MOD	9773	11 957	12 142	33 405
27 May				
Observations	297	473	464	2489
NVD-RLT	4169	5906	6379	36 312
THOMP	4834	8363	6465	51 236
NVD-MOD	5073	10 316	7130	48 188
Total				
Observations	1457	2776	2332	15 337
NVD-RLT	39 674	53 908	65 626	277 079
THOMP	45 054	59 472	62 971	338 042
NVD-MOD	46 692	60 397	64 760	317 579

both THOMP and NVD-MOD have an ensemble mean skill close to 0.3, compared to 0.2 on NVD-RLT. It is also apparent from this plot that much of the improvement is due to an increase in POD. Corresponding skill scores for CTP are similar, with the improvement of THOMP and NVD-MOD over NVD-RLT extending until the end of the 180-min (3 h) forecast (Figs. 15c–f). The performance diagram for CTP at the 60-min forecast time is also similar to the one for TB_{107}, with NVD-MOD being the best performer. Finally, skill scores for CWP are similar between all experiments for all forecast times (Figs. 15g–i).

Much of the improvement in skill arises from the reduction in object size in THOMP and NVD-MOD compared to NVD-RLT. These smaller object sizes
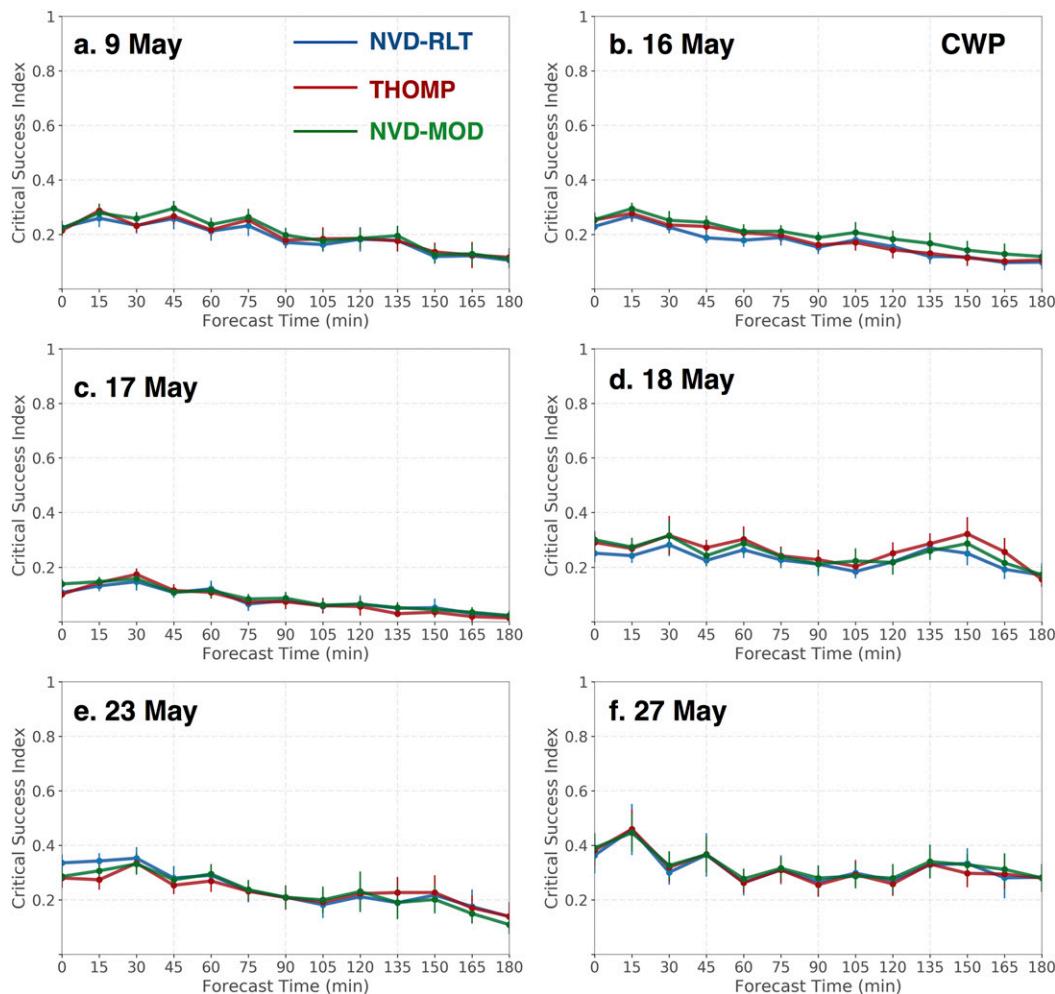
FIG. 14. As in Fig. 11, but for CWP objects.

better reflect the observed objects present in the *GOES-13* imagery. To illustrate these differences, Fig. 16 shows the median object area for TB_{107}, CTP, and CWP as a function of forecast time for the six-case sample split into matched and false alarm objects. For TB_{107} and CTP, matched objects are generally far larger than false alarm objects, indicating that the smallest objects in the model are not being matched to observations (Fig. 16a). Conversely, many of the larger objects are matched, resulting in much higher median area values (Fig. 16b). The NVD-RLT experiment produces the largest TB_{107} and CTP matched objects while those generated by THOMP and NVD-MOD are much smaller. Similar patterns are present for false alarm objects where NVD-RLT generates larger median values. Finally, there is somewhat less variability in CWP object size between experiments though NVD-MOD generally produces the smallest median values (Fig. 16c). As with TB_{107} and CTP, false alarm objects are smaller than matched objects. The very large satellite object sizes

generated by NVD-RLT are consistent with qualitative observations made during the 2017 HWT experiment, which provided the initial motivation for this work.

d. Reflectivity object skill

While the modifications to the NVD microphysics had a substantial positive impact when comparing satellite objects, it is also important to verify that these changes did not adversely impact the reflectivity skill given its importance in severe weather forecasting. Performance diagrams for $t = 60$, 120, and 150 min forecasts show that overall the skill (CSI) is similar for all experiments (Fig. 17). Note that 180-min performance diagrams are not shown since reflectivity uses a ±20-min window for object matching, which prevents full statistics from being computed after the 160-min forecast. For all forecast times, NVD-RLT generates the lowest POD, but with the advantage of also producing a lowest FAR (high success ratio). The THOMP and NVD-RLT experiments generate somewhat higher POD and FAR
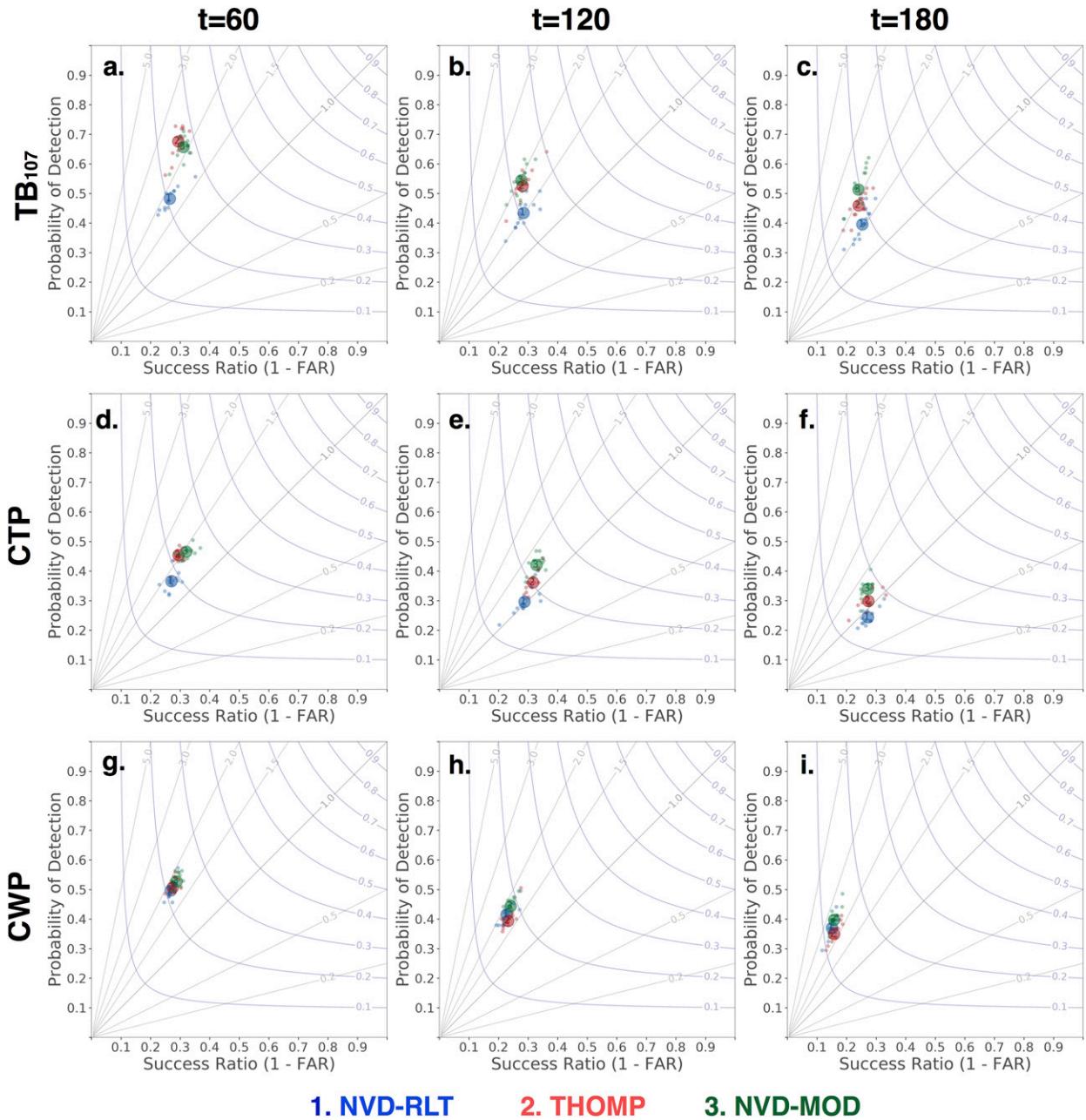
FIG. 15. As in Fig. 12, but showing the performance of all cases for each experiment at (left) 60-, (center) 120-, and (right) 180-min forecast times for (a)–(c) TB$_{107}$, (d)–(f) CTP, and (g)–(i) CWP.

values. As a result, the ensemble mean CSI changes little from case to case. By the 150-min forecast, individual member skill levels from all experiments overlap one another, indicating that the differences in reflectivity forecasts by this time are negligible. Thus, the impact of the microphysics modifications, which are large for cloud properties, do not have a negative impact on reflectivity forecasts, which is important for high-impact weather forecasts. Verification against other parameters such as

2–5-km updraft helicity was also conducted and while NVD-MOD performed somewhat worse than NVD-RLT, out 90 min, it was very similar afterward with the overall skill difference being small (not shown).

6. Conclusions

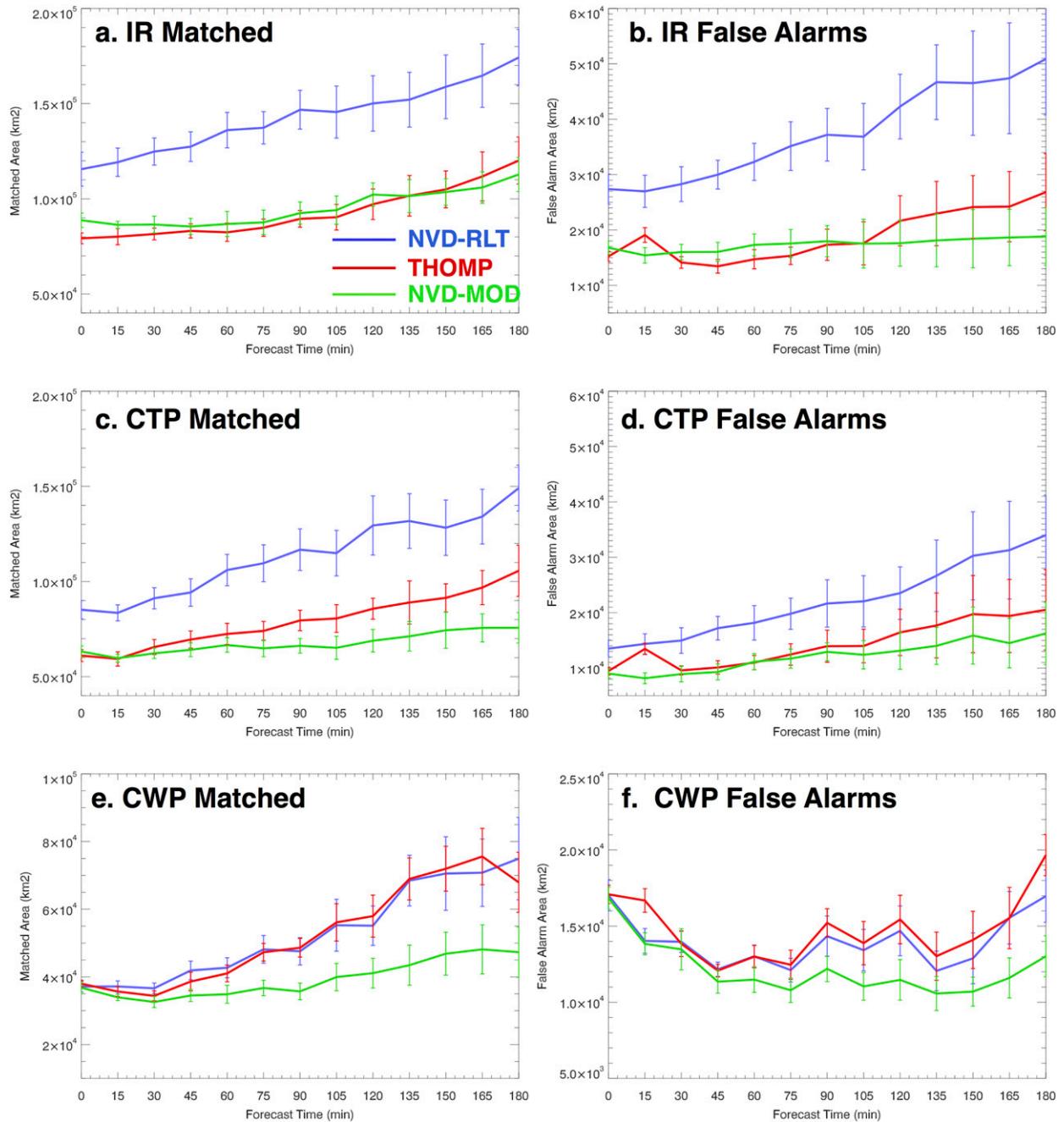Object-based verification of simulated satellite products against *GOES-13* observations showed that the

FIG. 16. Ensemble mean (left) matched and (right) false alarm object size as a function of forecast time for (a),(b) TB_{107}, (c),(d) CTP, and (e),(f) CWP objects averaged over all cases for each experiment. Vertical lines indicate the standard deviation of CSI calculated among all forecast members.

choice of the cloud microphysics scheme used within the model can have a large impact on the skill of these products. The model configuration employed during the 2017 HWT used the NVD double-moment cloud microphysics scheme developed by NSSL, which generated upper-level clouds that often had a much greater areal coverage than those present in observations. Since the

NVD scheme had never been validated using satellite data until this point, this large cloud bias had gone unnoticed. Comparisons of the real-time configuration with an identical configuration, but now using the Thompson microphysics scheme, showed that the latter generated more realistic upper-level cloud coverage. Thus, an effort was made to modify the NVD scheme to
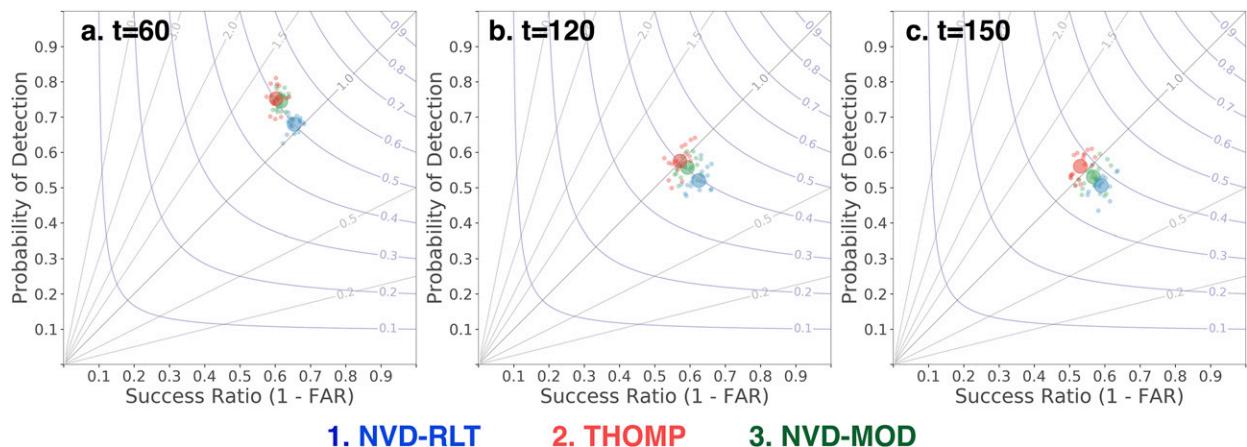
FIG. 17. Radar reflectivity performance diagrams for (a) 60-, (b) 120-, and (c) 150-min forecasts aggregated over all cases for each experiment.

improve the upper-level cloud properties while maintaining skill levels for precipitation. The modified NVD scheme was consistently more skillful when comparing TB_{107} and CTP against observations, while only a marginal improvement was observed for CWP. While both TB_{107} and CTP are essentially measures of cloud-top characteristics, it is noteworthy that both a raw radiances and cloud-top retrievals give similar results. Corresponding object-based radar reflectivity skill was similar between NVD-RLT and NVD-MOD experiments, indicating that the much larger changes to the upper-level cloud properties did not have any adverse impacts on the forecast reflectivity and precipitation.

The improvement in simulated TB_{107} and CTP skill occurs primarily through a reduction in object size made possible by the changes to the fall speed and collection efficiency variables. When assimilating cloudy observations, especially cloudy radiances, correctly analyzing the location and extent of the cloud cover is important since a drastic mischaracterization of cloud cover in the model will limit the impact of assimilating true observations, potentially reducing the overall model skill. This work has shown that modest adjustments to cloud microphysics scheme parameters can make a nontrivial difference and that both cloud and precipitation features should be considered when validating future adjustments to advance cloud microphysics schemes. Ongoing research will analyze the impact of microphysics changes on other variables, such as strong winds and accumulated precipitation using similar object-based verification approaches.

Continued object-based satellite verification will continue with the 2018 and future versions of the NEWS-e in order to test how future changes to the model configuration impact the cloud analysis and to make modifications as necessary. By verifying against both satellite and radar data,

we can tune the model configuration so that both the cloud and precipitation fields benefit. Higher spatial and temporal resolution data from *GOES-16* and *GOES-17* will be used as truth data and should enable even better assessment of NEWS-e skill, especially as it transitions to a higher horizontal resolution in the near future. Also, the increased spectral resolution from the additional infrared channel will allow for more accurate retrievals of cloud properties and improved sampling of the near-storm environment.

REFERENCES

Alexander, C. R., and Coauthors, 2018: Development of the High-Resolution Rapid Refresh Ensemble (HRRRE). *22nd Conf. on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface*, Austin, TX, Amer. Meteor. Soc., 11.3, https://ams.confex.com/ams/98Annual/webprogram/Paper335526.html.

Anderson, J. L., and N. Collins, 2007: Scalable implementations of ensemble filter algorithms for data assimilation. *J. Atmos.*

Oceanic Technol., **24**, 1452–1463, https://doi.org/10.1175/JTECH2049.1.

——, T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The Data Assimilation Research Testbed: A community data assimilation facility. *Bull. Amer. Meteor. Soc.*, **90**, 1283–1296, https://doi.org/10.1175/2009BAMS2618.1.

Benjamin, S. G., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, https://doi.org/10.1175/MWR-D-15-0242.1.

Brown, B. R., M. M. Bell, and A. J. Frambach, 2016: Validation of simulated hurricane drop size distributions using polarimetric radar. *Geophys. Res. Lett.*, **43**, 910–917, https://doi.org/10.1002/2015GL067278.

——, ——, and G. Thompson, 2017: Improvements to the snow melting process in a partially double moment microphysics parameterization. *J. Adv. Model. Earth Syst.*, **9**, 1150–1166, https://doi.org/10.1002/2016MS000892.

Cai, H., and R. E. Dumais, 2015: Object-based evaluation of a numerical weather prediction model's performance through storm characteristic analysis. *Wea. Forecasting*, **30**, 1451–1468, https://doi.org/10.1175/WAF-D-15-0008.1.

Chaboureau, J.-P., and J.-P. Pinty, 2006: Validation of a cirrus parameterization with Meteosat Second Generation observations. *Geophys. Res. Lett.*, **33**, L03815, https://doi.org/10.1029/2005GL024725.

Chevallier, F., and G. Kelly, 2002: Model clouds as seen from space: Comparison with geostationary imagery in the 11-μm window channel. *Mon. Wea. Rev.*, **130**, 712–722, https://doi.org/10.1175/1520-0493(2002)130<0712:MCASFS>2.0.CO;2.

——, P. Bauer, G. Kelly, C. Jakob, and T. McNally, 2001: Model clouds over oceans as seem from space: Comparison with HIRS/2 and MSU radiances. *J. Climate*, **14**, 4216–4229, https://doi.org/10.1175/1520-0442(2001)014<4216:MCOOAS>2.0.CO;2.

Choate, J. J., A. J. Clark, P. L. Heinselman, D. A. Imy, and P. S. Skinner, 2018: First demonstration of the NSSL Experimental Warn-on-Forecast System as part of the 2017 Spring Experiment. *Eighth Conf. on Transition of Research to Operations*, Austin, TX, Amer. Meteor. Soc., 1194, https://ams.confex.com/ams/98Annual/webprogram/Paper335289.html.

Cintineo, J. L., M. J. Pavolonis, J. M. Sieglaff, and A. K. Heidinger, 2013: Evolution of severe and nonsevere convection inferred from GOES-derived cloud properties. *J. Appl. Meteor. Climatol.*, **52**, 2009–2023, https://doi.org/10.1175/JAMC-D-12-0330.1.

Cintineo, R., J. Otkin, M. Xue, and F. Kong, 2014: Evaluating the performance of planetary boundary layer and cloud microphysical parameterization schemes in convection-permitting ensemble forecasts using synthetic *GOES-13* satellite observations. *Mon. Wea. Rev.*, **142**, 163–182, https://doi.org/10.1175/MWR-D-13-00143.1.

Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367–374, https://doi.org/10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2.

Davis, C. A., B. G. Brown, and R. G. Bullock, 2006a: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, **134**, 1772–1784, https://doi.org/10.1175/MWR3145.1.

——, ——, and ——, 2006b: Object-based verification of precipitation forecasts. Part II: Application to convective rain systems. *Mon. Wea. Rev.*, **134**, 1785–1795, https://doi.org/10.1175/MWR3146.1.

Dawson, D. T., II, L. J. Wicker, E. R. Mansell, and R. L. Tanamachi, 2012: Impact of the environmental low-level wind profile on ensemble forecasts of the 4 May 2007 Greensburg, Kansas, tornadic storm and associated mesocyclones. *Mon. Wea. Rev.*, **140**, 696–716, https://doi.org/10.1175/MWR-D-11-00008.1.

Dong, X., P. Minnis, B. Xi, S. Sun-Mack, and Y. Chen, 2008: Comparison of CERES-MODIS stratus cloud properties with ground-based measurements at the DOE ARM Southern Great Plains site. *J. Geophys. Res.*, **113**, D03204, https://doi.org/10.1029/2007JD008438.

Ebert, E. E., and W. A. Gallus Jr., 2009: Toward better understanding of the contiguous rain area (CRA) method for spatial forecast verification. *Wea. Forecasting*, **24**, 1401–1415, https://doi.org/10.1175/2009WAF2222252.1.

Ferrier, B. S., 1994: A double-moment multiple-phase four-class bulk ice scheme. Part I: Description. *J. Atmos. Sci.*, **51**, 249–280, https://doi.org/10.1175/1520-0469(1994)051<0249:ADMMPF>2.0.CO;2.

Gallo, B. T., and Coauthors, 2017: Breaking new ground in severe weather prediction: The 2015 NOAA/Hazardous Weather Testbed Spring Forecasting Experiment. *Wea. Forecasting*, **32**, 1541–1568, https://doi.org/10.1175/WAF-D-16-0178.1.

Gilleland, E., D. Ahijevych, B. Brown, and E. Ebert, 2009: Intercomparison of spatial forecast verification methods. *Wea. Forecasting*, **24**, 1416–1430, https://doi.org/10.1175/2009WAF2222269.1.

Grasso, L. D., and T. J. Greenwald, 2004: Analysis of 10.7-μm brightness temperatures of a simulated thunderstorm with two-moment microphysics. *Mon. Wea. Rev.*, **132**, 815–825, https://doi.org/10.1175/1520-0493(2004)132<0815:AOMBTO>2.0.CO;2.

——, and D. Lindsey, 2011: An example of the use of synthetic 3.9 μm *GOES-12* imagery for two-moment microphysical evaluation. *Int. J. Remote Sens.*, **32**, 2337–2350, https://doi.org/10.1080/01431161003698294.

——, M. Sengupta, J. F. Dostalek, R. Brummer, and M. DeMaria, 2008: Synthetic satellite imagery for current and future environmental satellites. *Int. J. Remote Sens.*, **29**, 4373–4384, https://doi.org/10.1080/01431160801891820.

——, ——, and M. DeMaria, 2010: Comparison between observed and synthetic 6.5 and 10.7 μm *GOES-12* imagery of thunderstorms that occurred on 8 May 2003. *Int. J. Remote Sens.*, **31**, 647–663, https://doi.org/10.1080/01431160902894483.

——, D. T. Lindsey, K.-S. Sunny-Lim, A. Clark, D. Bikos, and S. R. Dembek, 2014: Evaluation of and suggested improvements to the WSM6 microphysics in WRF-ARW using synthetic and observed *GOES-13* imagery. *Mon. Wea. Rev.*, **142**, 3635–3649, https://doi.org/10.1175/MWR-D-14-00005.1.

Griffin, S. M., J. A. Otkin, C. M. Rozoff, J. M. Sieglaff, L. M. Cronce, and C. R. Alexander, 2017a: Methods for comparing simulated and observed satellite infrared brightness temperatures and what do they tell us? *Wea. Forecasting*, **32**, 5–25, https://doi.org/10.1175/WAF-D-16-0098.1.

——, ——, ——, ——, ——, ——, T. L. Jensen, and J. K. Wolff, 2017b: Seasonal analysis of cloud objects in the High-Resolution Rapid Refresh (HRRR) model using object-based verification. *J. Appl. Meteor. Climatol.*, **56**, 2317–2334, https://doi.org/10.1175/JAMC-D-17-0004.1.

Han, Y., F. Weng, Q. Liu, and P. van Delst, 2007: A fast radiative transfer model for SSMIS upper atmosphere sounding channels. *J. Geophys. Res.*, **112**, D11121, https://doi.org/10.1029/2006JD008208.

Jones, T. A., D. J. Stensrud, L. Wicker, P. Minnis, and R. Palikonda, 2015: Simultaneous radar and satellite data storm-scale assimilation using an ensemble Kalman filter approach for 24 May 2011. *Mon. Wea. Rev.*, **143**, 165–194, https://doi.org/10.1175/MWR-D-14-00180.1.

——, K. Knopfmeier, D. Wheatley, G. Creager, P. Minnis, and R. Palikonda, 2016: Storm-scale data assimilation and ensemble forecasting with the NSSL Experimental Warn-on-Forecast System. Part I: Combined radar and satellite assimilation. *Wea. Forecasting*, **31**, 297–327, https://doi.org/10.1175/WAF-D-15-0107.1.

——, X. Wang, P. Skinner, A. Johnson, and Y. Wang, 2018: Assimilation of *GOES-13* Imager clear-sky water vapor (6.5 μm) radiances into a Warn-on-Forecast system. *Mon. Wea. Rev.*, **146**, 1077–1107, https://doi.org/10.1175/MWR-D-17-0280.1.

Jung, Y., M. Xue, and M. Tong, 2012: Ensemble Kalman filter analyses of the 29–30 May 2004 Oklahoma tornadic thunderstorm using one- and two-moment bulk microphysics schemes, with verification against polarimetric radar data. *Mon. Wea. Rev.*, **140**, 1457–1475, https://doi.org/10.1175/MWR-D-11-00032.1.

Keil, C., A. Tafferner, H. Mannstein, and U. Schättler, 2003: Evaluating high-resolution model forecasts of European winter storms by use of satellite and radar observations. *Wea. Forecasting*, **18**, 732–747, https://doi.org/10.1175/1520-0434(2003)018<0732:EHMFOE>2.0.CO;2.

Lindsey, D., D. Bikos, and L. Grasso, 2018: Using the *GOES-16* split window difference to detect a boundary prior to cloud formation. *Bull. Amer. Meteor. Soc.*, **99**, 1541–1544, https://doi.org/10.1175/BAMS-D-17-0141.1.

Line, W. E., T. J. Schmit, D. T. Lindsey, and S. J. Goodman, 2016: Use of geostationary Super Rapid Scan satellite imagery by the Storm Prediction Center. *Wea. Forecasting*, **31**, 483–494, https://doi.org/10.1175/WAF-D-15-0135.1.

Liu, C., and M. W. Moncrieff, 2007: Sensitivity of cloud resolving simulations of warm-season convection to cloud microphysics parameterizations. *Mon. Wea. Rev.*, **135**, 2854–2868, https://doi.org/10.1175/MWR3437.1.

Mace, G. G., T. P. Ackerman, P. Minnis, and D. F. Young, 1998: Cirrus layer microphysical properties derived from surface-based millimeter radar and infrared interferometer data. *J. Geophys. Res.*, **103**, 23 207–23 216, https://doi.org/10.1029/98JD02117.

——, Y. Zhang, S. Platnick, M. D. King, P. Minnis, and P. Yang, 2005: Evaluation of cirrus cloud properties from MODIS radiances using cloud properties derived from ground-based data collected at the ARM SGP site. *J. Appl. Meteor.*, **44**, 221–240, https://doi.org/10.1175/JAM2193.1.

Mansell, E. R., C. L. Ziegler, and E. C. Bruning, 2010: Simulated electrification of a small thunderstorm with two-moment bulk microphysics. *J. Atmos. Sci.*, **67**, 171–194, https://doi.org/10.1175/2009JAS2965.1.

Matsui, T., and Coauthors, 2014: Introducing multisensor satellite radiance-based evaluation for regional Earth system modeling. *J. Geophys. Res. Atmos.*, **119**, 8450–8475, https://doi.org/10.1002/2013JD021424.

Mecikalski, J. R., and K. M. Bedka, 2006: Forecasting convective initiation by monitoring the evolution of moving cumulus in daytime GOES imagery. *Mon. Wea. Rev.*, **134**, 49–78, https://doi.org/10.1175/MWR3062.1.

Milbrandt, J. A., and M. K. Yau, 2005a: A multimoment bulk microphysics parameterization. Part I: Analysis of the role of the spectral shape parameter. *J. Atmos. Sci.*, **62**, 3051–3064, https://doi.org/10.1175/JAS3534.1.

——, and ——, 2005b: A multimoment bulk microphysics parameterization. Part II: A proposed three-moment closure and scheme description. *J. Atmos. Sci.*, **62**, 3065–3081, https://doi.org/10.1175/JAS3535.1.

Minnis, P., and Coauthors, 2008a: Near-real time cloud retrievals from operational and research meteorological satellites. *Proc. SPIE*, **7108**, 710703, https://doi.org/10.1117/12.800344.

——, C. R. Yost, S. Sun-Mack, and Y. Chen, 2008b: Estimating the physical top altitude of optically thick ice clouds from thermal infrared satellite observations using CALIPSO data. *Geophys. Res. Lett.*, **35**, L12801, https://doi.org/10.1029/2008GL033947.

——, and Coauthors, 2011: CERES Edition-2 cloud property retrievals using TRMM VIRS and Terra and Aqua MODIS data—Part I: Algorithms. *IEEE Trans. Geosci. Remote Sens.*, **49**, 4374–4400, https://doi.org/10.1109/TGRS.2011.2144601.

——, and Coauthors 2016: A consistent long-term cloud and clear-sky radiation property dataset from the Advanced Very High Resolution Radiometer (AVHRR). Climate Algorithm Theoretical Basis Document (C-ATBD), Climate Data Record Rep. CDRP-ATBD-0826 Rev. 1, NASA, 159 pp.

Morrison, H., G. Thompson, and V. Tatarskii, 2009: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Mon. Wea. Rev.*, **137**, 991–1007, https://doi.org/10.1175/2008MWR2556.1.

Otkin, J. A., and T. J. Greenwald, 2008: Comparison of WRF Model-simulated MODIS-derived cloud data. *Mon. Wea. Rev.*, **136**, 1957–1970, https://doi.org/10.1175/2007MWR2293.1.

——, ——, J. Sieglaff, and H.-L. Huang, 2009: Validation of a large-scale simulated brightness temperature dataset using SEVIRI satellite observations. *J. Appl. Meteor. Climatol.*, **48**, 1613–1626, https://doi.org/10.1175/2009JAMC2142.1.

Painemal, D., P. Minnis, J. K. Ayers, and L. O'Neill, 2012: GOES-10 microphysical retrievals in marine warm clouds: Multi-instrument validation and daytime cycle over the southeast Pacific. *J. Geophys. Res.*, **117**, D06203, https://doi.org/10.1029/2012JD017822.

——, T. Greenwald, M. Cadeddu, and P. Minnis, 2016: First extended validation of satellite microwave liquid water path with ship-based observations of marine low clouds. *Geophys. Res. Lett.*, **43**, 6563–6570, https://doi.org/10.1002/2016GL069061.

Pinto, J. O., J. A. Grim, and M. Steiner, 2015: Assessment of the High-Resolution Rapid Refresh model's ability to predict mesoscale convective systems using object-based evaluation. *Wea. Forecasting*, **30**, 892–913, https://doi.org/10.1175/WAF-D-14-00118.1.

Rienecker, M. M., and Coauthors, 2008: The GEOS-5 Data Assimilation System—Documentation of versions 5.0.1, 5.1.0, and 5.2.0. NASA Tech. Memo. NASA/TM–2008–104606, Vol. 27, 118 pp., https://ntrs.nasa.gov/search.jsp?R=20120011955.

Roebber, P. J., 2009: Visualizing multiple measures of forecast quality. *Wea. Forecasting*, **24**, 601–608, https://doi.org/10.1175/2008WAF2222159.1.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., https://dx.doi.org/10.5065/D68S4MVH.

Skinner, P. S., L. J. Wicker, D. M. Wheatley, and K. H. Knopfmeier, 2016: Application of two spatial verification methods to ensemble forecasts of low-level rotation. *Wea. Forecasting*, **31**, 713–735, https://doi.org/10.1175/WAF-D-15-0129.1.

——, and Coauthors, 2018: Object-based verification of a prototype Warn-on-Forecast System. *Wea. Forecasting*, **33**, 1225–1250, https://doi.org/10.1175/WAF-D-18-0020.1.

Smith, T. M., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) severe weather and aviation products: Initial operating

capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 1617–1630, https://doi.org/10.1175/BAMS-D-14-00173.1.

Smith, W. L., Jr., 2014: 4-D cloud properties from passive satellite data and applications to resolve the flight icing threat to aircraft. Ph.D. dissertation, University of Wisconsin–Madison, 165 pp.

Snook, N., and M. Xue, 2008: Effects of microphysical drop size distribution on tornadogenesis in supercell thunderstorms. *Geophys. Res. Lett.*, **35**, 851–854, https://doi.org/10.1029/2008GL035866.

Sobash, R. A., G. S. Romine, C. S. Schwartz, D. J. Gagne, and M. L. Weisman, 2016: Explicit forecasts of low-level rotation from convection-allowing models for next-day tornado prediction. *Wea. Forecasting*, **31**, 1591–1614, https://doi.org/10.1175/WAF-D-16-0073.1.

Stensrud, D. J., J.-W. Bao, and T. T. Warner, 2000: Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems. *Mon. Wea. Rev.*, **128**, 2077–2107, https://doi.org/10.1175/1520-0493(2000)128<2077:UICAMP>2.0.CO;2.

——, and Coauthors, 2009: Convective-scale Warn-on-Forecast system: A vision for 2020. *Bull. Amer. Meteor. Soc.*, **90**, 1487–1499, https://doi.org/10.1175/2009BAMS2795.1.

——, and Coauthors, 2013: Progress and challenges with Warn-on-Forecast. *Atmos. Res.*, **123**, 2–16, https://doi.org/10.1016/j.atmosres.2012.04.004.

Straka, J. M., and E. R. Mansell, 2005: A bulk microphysics parameterization with multiple ice precipitation categories. *J. Appl. Meteor.*, **44**, 445–466, https://doi.org/10.1175/JAM2211.1.

Sun-Mack, S., P. Minnis, Y. Chen, S. Kato, Y. Yi, S. Gibson, P. W. Heck, and D. Winker, 2014: Regional apparent boundary layer lapse rates determined from *CALIPSO* and MODIS data for cloud height determination. *J. Appl. Meteor. Climatol.*, **53**, 990–1011, https://doi.org/10.1175/JAMC-D-13-081.1.

Thompson, G., and T. Eidhammer, 2014: A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. *J. Atmos. Sci.*, **71**, 3636–3658, https://doi.org/10.1175/JAS-D-13-0305.1.

——, R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.*, **132**, 519–542, https://doi.org/10.1175/1520-0493(2004)132<0519:EFOWPU>2.0.CO;2.

——, P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5114, https://doi.org/10.1175/2008MWR2387.1.

——, M. Tewari, K. Ikeda, S. Tessendorf, C. Weeks, J. Otkin, and F. Kong, 2016: Explicitly-coupled cloud physics and radiation parameterizations and subsequent evaluation in WRF high-resolution convective forecasts. *Atmos. Res.*, **168**, 92–104, https://doi.org/10.1016/j.atmosres.2015.09.005.

Tian, J., X. Dong, B. Xi, P. Minnis, W. L. Smith Jr., S. Sun-Mack, M. Thiemann, and J. Wang, 2018: Comparisons of ice water path in deep convective systems among ground-based, GOES, and CERES-MODIS retrievals. *J. Geophys. Res.*, **123**, 1708–1723, https://doi.org/10.1002/2017JD027498.

Tong, M., and M. Xue, 2008: Simultaneous estimation of microphysical parameters and atmospheric state with radar data and ensemble square-root Kalman filter. Part I: Sensitivity analysis and parameter identifiability. *Mon. Wea. Rev.*, **136**, 1630–1648, https://doi.org/10.1175/2007MWR2070.1.

Tselioudis, G., and C. Jakob, 2002: Evaluation of midlatitude cloud properties in a weather and a climate model: Dependence on dynamic regime and spatial resolution. *J. Geophys. Res.*, **107**, 4781, https://doi.org/10.1029/2002JD002259.

Weng, F., 2007: Advances in radiative transfer modeling in support of satellite data assimilation. *J. Atmos. Sci.*, **64**, 3799–3807, https://doi.org/10.1175/2007JAS2112.1.

Wheatley, D. M., K. H. Knopfmeier, T. A. Jones, and G. J. Creager, 2015: Storm-scale data assimilation and ensemble forecasting with the NSSL Experimental Warn-on-Forecast System. Part I: Radar data experiments. *Wea. Forecasting*, **30**, 1795–1817, https://doi.org/10.1175/WAF-D-15-0043.1.

Wolff, J. K., M. Harrold, T. Fowler, J. H. Gotway, L. Nance, and B. G. Brown, 2014: Beyond the basics: Evaluating model-based precipitation forecasts using traditional, spatial, and object-based methods. *Wea. Forecasting*, **29**, 1451–1472, https://doi.org/10.1175/WAF-D-13-00135.1.

Xie, B., J. C. H. Fung, A. Chan, and A. Lau, 2012: Evaluation of nonlocal and local planetary boundary layer schemes in the WRF model. *J. Geophys. Res.*, **117**, D12103, https://doi.org/10.1029/2011JD017080.

Yussouf, N., and D. J. Stensrud, 2012: Comparison of single-parameter and multiparameter ensembles for assimilation of radar observations using the ensemble Kalman filter. *Mon. Wea. Rev.*, **140**, 562–586, https://doi.org/10.1175/MWR-D-10-05074.1.

——, E. R. Mansell, L. J. Wicker, D. M. Wheatley, and D. J. Stensrud, 2013: The ensemble Kalman filter analyses and forecasts of the 8 May 2003 Oklahoma City tornadic supercell storm using single- and double-moment microphysics schemes. *Mon. Wea. Rev.*, **141**, 3388–3412, https://doi.org/10.1175/MWR-D-12-00237.1.

Ziegler, C. L., 1985: Retrieval of thermal and microphysical variables in observed convective storms. Part I: Model development and preliminary testing. *J. Atmos. Sci.*, **42**, 1487–1509, https://doi.org/10.1175/1520-0469(1985)042<1487:ROTAMV>2.0.CO;2.